A CRYPTOGRAPHIC SERIES

56

ELEMENTARY CRYPTOGRAPHY AND CRYPTANALYSIS

	1	2	3	4	5
16	A	B	C	D	E
17	F	G	H	I	K
18	L	M	N	O	P
19	Q	R	S	T	U
20	V	W	X	Y	Z

Donald D. Millikin

om Aegean Park Press

MIL

005.82 MIL
MILLIKIN, DONALD D.
ELEMENTARY CRYPTOGRAPHY A
OBC -0033516 .000 ?bb?

ELEMENTARY CRYPTOGRAPHY AND CRYPTANALYSIS

מכללת אורט בראודה
ספריה
קרית המכללה 78 כרמיאל 21982
מס' פרים 33516-000-001
מס' מיון 005.82
MIL

DONALD D. MILLIKIN

This is a quality reproduction of a classic text written and used by Donald D. Millikin to introduce cryptography and cryptanalysis to college students at New York University in 1943.

ISBN: 0-89412-173-1 (soft cover)

ISBN: 0-89412-174-X (library bound)

AEGEAN PARK PRESS

P.O. Box 2837

Laguna Hills, California 92654

(714) 586-8811

Manufactured in The United States of America

CONTENTS

		Page
	INTRODUCTION	ii
	Historical Notes	v
	Methods of Secret Communication	vi
CHAPTER		
I	Preliminary Definitions.	1
II	The Written Language. Kinds of Alphabets. Keys: Literal and Numerical. Construction of Key-word Mixed Alphabets	4
III	Frequency Charts, Graphs and Tables.	14
IV	Rules for the Use of Codes and Ciphers. Hints and Suggestions. Some Mechanical Aids.	24
V	On the Analysis of Cryptograms. Methods of Analysis. Frequential Analysis Versus the "Probable Word."	28
VI	The Two Principal Classes of Ciphers. Distinguishing Between Transposition and Substitution	33
VII	Transposition Systems. Route Transpositions With Geometric Figures. Decrypting Geometric Route Transpositions. The "Rail-Fence Cipher."	37
VIII	Columnar Transposition Ciphers with Completely Filled Figures. The "Federal Route Cipher" of the "Civil War." Decrypting the Cipher	48
IX	Columnar Transposition Ciphers with Incompletely Filled Figures. Cryptanalysis of Incomplete Rectangles.	54
X	Simple Substitution. Monoalphabetic Systems. Decryptment, (Frequential Analysis). The "Probable Word." Solution by Comparison of Frequency Graphs. Solution by Completing the Plain Component, ("Running Down the Alphabet.")	59
XI	Monoalphabetic Ciphers Using Symbols and Numerals. Cryptanalysis of a Cipher Using Symbols. Systems Employing Numerals.	70

CHAPTER

		Page
XII	Miscellaneous Monoalphabetic Ciphers. Poly-Literal Systems. Sir Francis Bacon's Bi-Literal Cipher. The Vowel Cipher. A Tri-Literal System. Monoalphabetic Substitution with Variants. Homophones. Homophonic Systems	77
XIII	Polyalphabetic Systems. Periodic Ciphers and "Cyclic Phenomena." The Principle of Factoring	83
XIV	Square Tables, Discs and Slide Rules. The Vigenère Cipher. Porta's Table. The Beaufort Cipher. The Army Signal Disc. A Slide Rule. A Modified Vigenère	88
XV	Digraphic Systems. The Playfair. Construction of a Playfair Square by Means of the Knight's Move in Chess. Rules for Enciphering. Example of Encipherment. Decipherment. Peculiarities of the Playfair and Rules for Recognizing and Solving Cryptograms Enciphered by the Playfair System. The Four Square Matrix Cipher	98
XVI	Cryptograms in Foreign Languages. Concluding Remarks. Sir Francis Bacon's Rules for Ciphers. Fundamental Requirements for a Military Cipher.	106

APPENDICES

A	Glossary of Terms	108
B	Bibliography of Works in English	114
C	Language Data: English, French, German, Italian, and Spanish	115
D	Special Lists: English	126
E	Miscellaneous Data	129

INTRODUCTION

Secret communications were first mentioned by a military writer in "POLIORCETICUS," by AEneas the Tactician in the 4th Century B.C.

Any code or cipher that has ever been used may be used again, so that an acquaintance with their history is necessary. A thorough knowledge of both cryptography and cryptanalysis is an essential requirement for those engaged in either branch of the art. The importance of cryptanalysis, from a military standpoint, is evidenced by the statement of General Cartier, one of the foremost of the French experts, that during World War Number One, it "was indisputably superior to all other means for securing intelligence."

Training in cryptanalysis takes a long time. The late General Marcel Givierge, who at one time headed the French Code and Cipher Bureau in the last war, stated that years were required to develop adepts in this art. Instruction in cryptography and cryptanalysis go hand-in-hand. Each is necessary to the other.

Colonel Parker Hitt, in his "Manual for the Solution of Military Ciphers," says:

"SUCCESS in dealing with unknown ciphers is measured by these four things in the order named: - perseverance, careful methods of analysis, intuition, luck. The ability to at least read the language of the original text is very desirable but not essential."

"Cipher work will have little permanent attraction for one who expects results at once, without labor, for there is a vast amount of routine labor in the preparation of frequency tables, the re-arrangement of ciphers for examination, and the trial and fitting of letter to letter before the message begins to appear."

"Under intuition must be included a knowledge of the general situation--."

"As to luck, there is the old miner's proverb: "Gold is where you find it."

The motto of the cryptanalyst is "SUPPOSE THAT." Having made the supposition he must discover all the facts that he can obtain to confirm it.

Historical Notes

CRYPTOGRAPHY (from the Greek words, Kryptos - "secret" and Graphein - "to write") has a history reaching back considerably more than two thousand years. Some of the earliest examples recorded are found in the Bible in the Books of Jeremiah (XXV.-26, and LI.-41) and Isaiah (VII.-6) where SHESHACH and TABEAL are the cipher names used in place of BABEL and REMALIAH.

The Greeks and Romans employed ciphers. The "Scytala" of the Spartans was one of the earliest devices for sending military communications. Julius Caesar and the Emperor Augustus both used systems which shifted the alphabet.

From the time of the Caesars to the Renaissance period cryptography developed very slowly. During the next four hundred years of the thirteenth to sixteenth centuries advancement in the use of codes and ciphers was marked. It was then that many of the important works on the subject were written and some of the systems employed at that time form the basis of our modern code and cipher methods.

Mary Stuart, Queen of Scots, and Charles the First, lost their heads through the solving of secret messages by the cryptanalysts in the employ of Elizabeth and Cromwell. Dr. John Wallis, Savillian Professor of Mathematics at Oxford, deciphered the cryptograms found among the papers of Charles the First after the Battle of Naseby. Two other cryptanalysts of note were Francois Viete, the celebrated French mathematician, who solved the messages of the Holy League, and Antoine Rossignol, who broke the ciphers of the Huguenots during the reign of Louis XIII.

In the American Revolution, both the Colonists and the British made use of secret writing. Napoleon, too, had his corps of experts.

No sooner had the telegraph come into use, than commercial firms promptly devised code and cipher systems. The Civil War in the United States offered an opportunity for the extensive use of cryptography in military communications, while World War Number One and the employment of radio made even more necessary the use of codes and ciphers.

In addition to its applications in diplomatic, military and naval correspondence, cryptography has entered into many other phases of every day life. Codes and ciphers are used in commerce and industry for purposes of economy and secrecy. Spies and criminals employ them. In the field of literature acrostics and anagrams abound from early times to the present day. Edgar Allen Poe's "The Gold Bug," and his writings on the analysis of ciphers have delighted multitudes. Conan-Doyle's "Adventure of the Dancing Men" and Jules Verne's "Voyage to the Center of the Earth," provide two examples of ciphers in fiction. Secret messages appear quite frequently in modern detective stories. Modern shorthand is based partly on Timothy Bright's "The Arte of Shorte, Swifte and Secrete Writing," published in 1588. In the field of recreation we have acrostics, anagrams, rebuses and other forms of cryptography, while a number of newspapers and magazines present cryptograms to their readers for solution.

Methods of Secret Communication

Codes and Ciphers are the two main divisions of visible secret writing. Pictures and drawings provide two other means of secret communication. Invisible writing includes microscopic inscriptions and sympathetic inks. In olden times the Greeks made use of lemon juice and other liquids, which were invisible until heat was applied. Other inks are brought out by chemical and physical means. Microscopic inscriptions are written or engraved under a microscope or a powerful lens and a magnifying glass of high power is required to read them.

From the time the fall of troy was signalled to the mainland of Greece by

flares, or the burnished shields of soldiers flashed their messages, torches and other forms of secret communication by means of light have been used.

The smoke signals of the American Indians and of other tribes are well known. Flags and banners, as well as panels of cloth or other materials have carried their messages from very early times. Of course, the use of couriers, on foot or mounted, was one of the earliest methods of carrying secret communications. Carrier pigeons were employed by Julius Caesar in his Gallic Wars and are still used by armies today.

Signalling by sound is also of ancient origin. The drums of the African tribes, the shots of cannon and other fire-arms, and bells, have all carried secret information. Within the past fifteen years the "scrambled speech" for transoceanic telephoning has given a measure of secrecy between persons separated by the seas.

Experiments have been tried, with varying success, in the use of the ultra-violet and infra-red rays.

The literature of cryptography is extensive, but the majority of the works are in foreign languages and many publications are rare. The scarcity of treatises in English is marked. Texts on cryptanalysis are even more rare than those dealing with cryptography. Most of them are government publications and not available to the general reader. A brief bibliography will be found in Appendix B.

Throughout this text the word, ENEMY, will be used to designate any persons not authorized to have access to the messages.

Chapter I

PRELIMINARY DEFINITIONS

A definite nomenclature and certain definitions, which have become standard through usage, are necessary to an understanding of the subject. Cryptography is the science which deals with all the means and methods for converting an intelligible message into unintelligible or secret form and for reconverting the secret form into the intelligible message by a direct reversal of the steps used in the original process. The words of the intelligible message constitute the plain text, (also called "clear text"), and the characters of the secret form of the message constitute the secret text or cryptogram. Cryptanalysis is the art of resolving cryptograms into their intelligible texts without having possession of the system or key employed. In solving unknown cryptograms, use is made of the characteristics of the language suspected and every scrap of information from all available sources concerning the circumstances under which the messages were sent and obtained is utilized.

Codes and ciphers are the two principal methods of secret communications, although a code is really a form of substitution cipher. The distinction is more or less arbitrary and the definitions have been adopted for the sake of convenience. A code handles elements of varying lengths - letters, syllables, words, phrases, and even whole sentences, while a cipher usually deals with the individual letters or pairs of letters of a message. The use of code generally necessitates the possession by each correspondent of identical copies of code books or code lists. The code groups may be composed of either letters or figures. Codes may be one- or two-part. One-part codes have the plain text meanings arranged in alphabetical order, with their code groups - either in alphabetical or numerical sequence - beside them. This type of code can be used for either encoding or decoding. Encoding, or simply coding, is the conversion of plain text into secret form.

Decoding is the reverse process. A two-part code, or random code, consists of two sections, one for encoding and one for decoding. The encoding section has the meanings listed in alphabetical order, but the code group equivalents of the expressions are in random sequence as if they had been mixed up and drawn out of a hat. In the decoding section the code groups are arranged in alphabetical or numerical order and the meanings appear in random or non-alphabetical sequence.

Ciphers are divided into two main classes, transposition and substitution. In a transposition cipher the individual letters, pairs of letters, or words, of the plain text have simply been re-arranged in a different order. Their positions with relation to each other have been changed. Writing a message backwards is a simple type of transposition. "GO AT ONCE" would appear as "ECNO TA OG." In a substitution cipher the elements of the plain language message - letters, pairs of letters, etc. - keep their usual relative positions, but they are replaced in the cipher text by other letters, or figures, or symbols. "GO AT ONCE" might be written as XP SO PEVY, or 11 24 09 32 24 10 17 06, or $& *# &/%(.

Most cipher systems consist of two parts; a general method which is fixed, and a key - variable at the will of the correspondents - that controls the operation of the basic method. Keys are either literal, consisting of words, phrases or sentences; or numerical, composed of figures.

Transposition ciphers are of various types - reversed writing, anagrams, route ciphers, "grid" or grille systems, and so on. Many of them are based on geometric patterns. Transpositions may be double or single.

Substitution systems are far more numerous and much more practical. Among the many substitution methods may be mentioned string ciphers; square tables, slides, discs or "wheel" ciphers; and symbols, including the so-called "tic-tac-toe" or angle-writing.

Acrostics, stencil or "window" systems, and picture ciphers almost form a class by themselves.

Transposition and substitution may be used in combination. A cryptogram produced by a substitution system may undergo further encipherment by the same or another substitution method, in which case double substitution is involved. Messages already in code may be given further security by encipherment through transposition or substitution and the resulting text is called enciphered code.

Substitution systems are monoalphabetic if a single cipher alphabet is used, polyalphabetic with two or more cipher alphabets. When the substitution involves one letter at a time the cipher is monographic. In a polygraphic system the substitution is by pairs of letters or larger regular groups. A well-known example is the digraphic cipher called the PLAYFAIR, in which each pair of letters in the plain text is enciphered by a pair of letters in the cryptogram.

Mechanical devices, usually hand-operated, are called "cryptographs." This noun should not be confused with the verb, cryptograph, which means to convert the plain text into the cryptogram. Cipher machines are either mechanically or electrically operated, or both, and many of them are extremely complicated.

A more complete list of definitions will be found in The Glossary, Appendix A.

Chapter II

THE WRITTEN LANGUAGE

Most languages, with few exceptions and those notably some of the Oriental tongues, are written by means of alphabets consisting of characters which we call letters and which represent individual sounds or combinations of sounds in the language. Thus, English has an alphabet of twenty-six letters, usually written from left to right

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.

The number of letters in an alphabet varies with the particular language. Hawaiian has only twelve, Greek and Latin twenty-four letters, and Russian thirty-five.

The word "cat" is written in English with the three letters c, a, and t, representing three sounds corresponding to the letters. We are so accustomed to our alphabet, that when we see the word "cat" we recognize it immediately and understand its meaning. Our mind at once pictures the pussy. But suppose that, instead of the letters c, a, and t, it has been our custom to use the letters o, z and u, to represent the same sounds. The word "ozu" would still be pronounced "cat" and we would recognize it as meaning the household pet. It is simply a matter of becoming used to the representation of sounds by an agreed-upon set of characters. There has been no change in the sounds themselves nor in the pronunciation of the word nor in the meaning conveyed. A different set of characters, (letters), has been employed.

The individual letters of an alphabet do not necessarily represent all the possible sound combinations in the language. In the word "church," it takes two letters c and h, in combination, to give the sounds of CH at the beginning and the end of the word. This combination of two letters is called a digraph. The "sh" in the word "show" is another; so is the "ti" in "action;" and the "er" in "farmer." When three letters combine they form a trigraph. The words bigram and trigram are

sometimes used in place of digraph and trigraph.

In any alphabet certain letters combine more readily with others to form digraphs, trigraphs, syllables and words in the language to which the alphabet belongs. In English the vowels, a, e, i, o, and u, join readily with all of the consonants. L, n, r, s, and t, and to a lesser degree c, d, f, h, m, and p, are found in frequent combinations with other letters. Each language, of course, has its own characteristics. Certain letters can precede others but are not often found following the same ones. The reverse is also true.

KINDS OF ALPHABETS

A normal alphabet is one that is written in the usual sequence and direction, according to the custom of the language. Most of those we are familiar with are written from left to right, but Hebrew, for example, is written from right to left. This is the normal alphabet in English:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.

Besides the normal alphabet, there is also the cipher alphabet of the secret text. Cipher alphabets are of two kinds, standard and mixed. The cipher alphabets may commence with any one of the twenty-six letters.

A standard alphabet runs in sequence from A to Z and may be written from left to right, in which case it is called a direct standard alphabet; or, it may be written in the reverse direction from right to left, forming a reversed standard alphabet. The normal alphabet is also a direct standard alphabet. Another example is given below, starting with the fourth letter:

D E F G H I J K L M N O P Q R S T U V W X Y Z A B C.

Two specimens of reversed standard alphabets are:

Z Y X W V U T S R Q P O N M L K J I H G F E D C B A,

P O N M L K J I H G F E D C B A Z Y X W V U T S R Q.

Mixed alphabets are of two types, systematically mixed and random mixed. Systematically mixed alphabets are usually formed by means of a key and there are a number of ways of constructing such alphabets. These will be treated under the

subject of "Keys". One example, where the key-word is written first followed by the other letters of the alphabet not in the key-word, is this:

S T E N O G R A P H I C B D F J K L M Q U V W X Y Z.

A random mixed alphabet would be:

W N H P U Y S X Q Z R O T G L E M A J B F K D I V C.

All cipher alphabets may be classified as enciphering alphabets and deciphering alphabets, arranged for convenience in enciphering and deciphering cryptograms. Each is composed of two parts, the plain text component and the cipher component. The latter is sometimes called the "crypto" component. With the random-mixed alphabet shown above, they would be:

Enciphering Alphabet

Plain component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.

Cipher component - W N H P U Y S X Q Z R O T G L E M A J B F K D I V C.

Deciphering Alphabet

Cipher component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.

Plain component - R T Z W P U N C X S V O Q B L D I K G M E Y A H F J.

In the enciphering alphabet the letters of the plain text are found written in normal sequence in the plain component with their equivalents in the cipher text just below them in the cipher component. Thus the word ENEMY would be enciphered UGUTV. In the deciphering alphabet the letters of the cryptogram are found in the normal sequence of the cipher component, with their equivalent values in the plain text underneath them in the plain component in random sequence.

It is customary in cryptography and cryptanalysis to designate a letter of the plain text by the lower-case letter p suffixed to it thus A_p. In a similar manner a letter of the cipher text is designated by the lower-case suffix c, thus R_c. This system is useful in cryptographic equations. The fact that A in the plain text is represented by R in the cryptogram would be shown in this way:- A_p equals R_c, or the reverse - R_c equals A_p.

A reciprocal alphabet is one in which the values of the letters of the alphabet are all reciprocal in pairs. Below is one example.

Enciphering Alphabet

Plain - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher - R K V Q Z U S O P W B X T Y H I D A G M F C J L N E .

Deciphering Alphabet

Cipher - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Plain - R K V Q Z U S O P W B X T Y H I D A G M F C J L N E .

Note that the enciphering and deciphering alphabets are identical and that it is not necessary to write out both. Either one can be used to encipher or to decipher. All reversed standard alphabets are reciprocal, in so far as each is concerned, but not to the others. That is clear from the two alphabets shown on page 5, as arranged below.

Plain - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher - Z Y X W V U T S R Q P O N M L K J I H G F E D C B A-1st Alphabet,

Cipher - P O N M L K J I H G F E D C B A Z Y X W V U T S R Q-2nd Alphabet.

Keys

Very often it becomes necessary in cryptographic operations to make use of literal and numerical keys. A literal key consists of the letters of a word, phrase, sentence, or even a whole paragraph. In choosing a literal key the following factors are important.

a. It should be easily remembered. LET US GO HOME is better than CONSTITUTIONAL.

b. It should consist of one or more simple words that can be spelled only one way. REINFORCE or REENFORCE and LABOR or LABOUR are undesirable.

c. It should contain as many different letters as possible. BANKRUPTCY and LOGARITHMS are preferable to ELEMENTARY and INDICATION.

d. It should have no association with, or relationship to, the special

situation or circumstances. If action is taking place in the vicinity of NEW YORK CITY, such words as BOWLING GREEN, RICHMOND, NEWARK, should be eliminated.

Numerical Keys

The remembering of a long string of numbers is difficult. To avoid this and to eliminate the necessity of keeping a written memorandum of the numbers, very simple methods have been devised. One is to assign numerical values to the letters of a key-word or a key-phrase according to the numerical position of the letters in the normal alphabet. Suppose the key-word QUESTION is selected. The letters A, B, C, and D, do not appear in the key-word, but E does; therefore under the letter E is placed the digit 1. The next letter of the alphabet occurring in the key-word is I, so below it is written the figure 2:-

Q U E S T I O N

1 2

Then comes N, with 3 underneath, and so on.

Q U E S T I O N

5 8 1 6 7 2 4 3

The derived numerical key is therefore: 5-8-1-6-7-2-4-3.

Other Examples

1. Key-word - S T E N O G R A P H I C

   Derived Numerical Key - 11 - 12 - 3 - 7 - 8 - 4 - 10 - 1 - 9 - 5 - 6 - 2

2. Key-word - F R E D E R I C K

   Derived Numerical Key - 5 - 8 - 3 - 2 - 4 - 9 - 6 - 1 - 7

3. Key-word - W E S T I N G H O U S E

   Derived Numerical Key - 12 - 1 - 8 - 10 - 5 - 6 - 3 - 4 - 7 - 11 - 9 - 2

The Key-word Method of

Constructing Systematically Mixed Alphabets

On page 6 is shown the systematically mixed alphabet obtained by writing the key-word STENOGRAPHIC, followed by the other letters of the alphabet not in the key-word. The enciphering and deciphering alphabets in this case are:

Enciphering Alphabet

Plain - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher - S T E N O G R A P H I C B D F J K L M Q U V W X Y Z .

Deciphering Alphabet

Cipher - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Plain - H M L N C O F J K P Q R S D E I T G A B U V W X Y Z .

Note the underlined sequences which show the weakness of the key-word chosen. The alphabets are not thoroughly mixed, especially UVWXYZ.

Let us try another key-word, HYDRAULICS.

Enciphering Alphabet

Plain - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher - H Y D R A U L I C S B E F G J K M N O P Q T V W X Z .

Deciphering Alphabet

Cipher - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Plain - E K I C L M N A H O P G Q R S T U D J V F W X Y B Z .

With the key-word STENOGRAPHIC, we have six letters at the end of our alphabet, all enciphered by themselves, UVWXYZ. Using HYDRAULICS, we have only one letter enciphered by itself - Z, in spite of the underlined sequences, which still constitute a weakness, though not vital as in the case of the other key.

Besides this method of constructing systematically-mixed cipher alphabets, four others will be given, all of them transposition processes.

Simple Columnar Transposition, with Key-word

Key-word - RADIOGRAMS:

R A D I O G M S

B C E F H J K L

N P Q T U V W X

Y Z

Each column is taken in order, giving the mixed sequence:

R B N Y A C P Z D E Q I F T O H U G J V M K W S L X

which yields the following enciphering alphabet:

Plain component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Cipher component - R B N Y A C P Z D E Q I F T O H U G J V M K W S L X

Numerical Key Columnar Transposition

Key-word - UNIFORMS:

Numerical Key - 8-4-2-1-5-6-3-7

U N I F O R M S

A B C D E G H J

K L P Q T V W X

Y Z

Mixed sequence is obtained by writing first the letters in the column with 1 above it, F D Q, then the letters in the column headed 2, I C P, continuing with the remaining columns according to the number at the top. The mixed sequence is:

F D Q I C P M H W N B L Z O E T R G V S J X U A K Y ,

and the enciphering alphabet:

Plain component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher component - F D Q I C P M H W N B L Z O E T R G V S J X U A K Y .

The Decimation Method

The decimation method consists, first, of an agreement upon a number; second, counting off the letters in a normal alphabet or in a systematically mixed alphabet.

The alphabet to be decimated is considered as a circle. As each letter is counted off, it is written in a new sequence and dropped as the count is continued. With the number 9 and the key-word mixed alphabet,

J O N E S B U R G A C D F H I K L M P Q T V W X Y Z ,

the procedure is as follows: count the first nine letters and place the number 1 below the 9th letter, G; then count another nine letters, reaching M, under which the number 2 is placed;

J O N E S B U R G A C D F H I K L M P Q T V W X Y Z .

1 2

The third count begins with the letter P and goes back to the start of the alphabet, the number 9 falling on the letter J; beginning the fourth count of nine on the letter O, we disregard G because we have already placed it as the first letter of our sequence G M J...

J O N E S B U R G A C D F H I K L M P Q T V W X Y Z

3 1 4 2 5

and we reach the letter C, under which is written the number 4; similarly, on the fifth count of nine M is disregarded and T becomes the fifth letter of the sequence, the entire sequence being:

G M J C T S K O H Z I E Q A N W P L V Y R B U X D F .

The enciphering alphabet is thus:

Plain component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z .

Cipher component - G M J C T S K O H Z I E Q A N W P L V Y R B U X D F .

A Double Transposition Method*

With this method a key-word of five letters is chosen and the derived numerical key found. The key-word is written in the first row of a square matrix consisting of twenty-five small squares or cells. The other letters of the alphabet that do not occur in the key-word are written in order in the remaining four rows, I and J being considered as the same letter. The derived numerical key is then written horizontally above the matrix and vertically at its left side. With the key-word GAMES and the derived numerical key 3-1-4-2-5, the figure will be:

3-1-4-2-5

3	G A M E S
1	B C D F H
4	I K L N O
2	P Q R T U
5	V W X Y Z

The digits of the numerical key at the top are the column indicators and those at the left indicate the rows, in the matrix. The column and row indicators are written in two-digit figures and the letters of the mixed sequence are written below the figures according to the positions of the letters in the matrix as determined by the column and row indicators. The first letter placed in the sequence is found in Column 1, Row 1. This is C, the second letter being in Column 1, Row 2; the third will be in Column 1, Row 3, and so on. The sixth letter is in Column 2, Row 1. The 25-letter sequence will be:

11 12 13 14 15 21 22 23 24 25 31 32 33 34 35 41 42 43 44 45 51 52 53 54 55

O Q A K W F T E N Y B P G I V D R M L X H U S O Z

We omitted the letter J in forming the matrix. J can now be placed either at the beginning or at the end of the 25 letter mixed sequence or in any other position agreed upon by the correspondents. The enciphering and deciphering alphabets are; with J at the end: -

---

*NOTE. This method and the decimation process are described by F. Delastelle in his "Traité Elémentaire de Cryptographie," Paris, 1902, pages 47 and 48.

Enciphering

Plain - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher - C Q A K W F T E N Y B P G I V D R M L X H U S O Z J .

Deciphering

Cipher - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Plain - C K A P H F M U N Z D S R I X L B Q W G V O E T J Y .

A variation of this method employs two different keys for determining the columns and rows, of which the following is an example. Key-word for the matrix and column indicators, GAMES, (numerical key 3-1-4-2-5); for the row indicators - UNITE and 5-3-2-4-1.

3-1-4-2-5

U	5	G	A	M	E	S
N	3	B	C	D	F	H
I	2	I	K	L	N	O
T	4	P	Q	R	T	U
E	1	V	W	X	Y	Z

The mixed sequence will be:

11 12 13 14 15 21 22 23 23 25 31 32 33 34 35 41 42 43 44 45 51 52 53 54 55

W K C Q A Y N F T E V I B P G X L D R M Z O H U S

Placing J at the beginning results in the following enciphering alphabet:

Plain - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z ,

Cipher - J W K C Q A Y N F T E V I B P G X L D R M Z O H U S .

These two double transpositions produce alphabets that are more thoroughly mixed than is the case with any of the other ways that we have described, except that of decimation.

In addition to their employment for the construction of systematically mixed cipher alphabets, literal and numerical keys have other uses which will be discussed as the particular applications in cipher operations are taken up.

Chapter III

FREQUENCY CHARTS, GRAPHS AND TABLES

If the consecutive letters of any considerable volume of English text are counted and the number of times each letter of the alphabet occurs is set down, the frequency of occurrence of each letter will closely approximate the number of times each letter will appear in any other English text of the same length. Variations from the normal frequency may be expected, depending upon the kind of text involved, such as literary, newspaper, technical, or telegraphic. As a rule, such variations are not great and the value of the normal frequencies as a guide is not impaired. It is customary to calculate on the basis of 10,000 letters, but for practical purposes it has been found that the frequencies in a text of 200 are usually adequate.

Listing the individual occurrences of the various letters of the alphabet yields a single frequency table. Appearances of pairs of letters are arranged in a table of digraphs, and the recurrences of groups of three are listed in a trigraphic frequency table. The table of digraphs is of great value when working with a cryptogram produced by a digraphic system, especially one of the Playfair type. The trigraphic table is perhaps the most important of all for it not only furnishes us with information about which letters precede and which letters follow each letter, but it can at the same time serve as a single frequency table and as a table of digraphs. It is valuable in determining which of the characters in a cryptogram are vowels and which are consonants. The trigraphic table is also designated by some authorities as a table of prefixes and suffixes. All three tables are used in the method of frequential analysis. Much of the work in solving cryptograms involves the preparation and use these tables. They are occasionally called frequency charts.

Frequency data on the English, French, German, Italian, and Spanish languages will be found in Appendix C.

There are several ways of preparing frequency tables. One of them is shown below, both in alphabetical order and in order of frequency, for English telegraphic text with a count of 200 letters (according to Hitt), Fig. 1.

(a)

In Alphabetic Order	
A- 11111 11111 11111 1	16
B- 111	3
C- 11111 1	6
D- 11111 111	8
E- 11111 11111 11111 11111 11111 1	26
F- 1111	4
G- 1111	4
H- 11111 111	8
I- 11111 11111 1111	14
J- 1	1
K- 11	2
L- 11111 111	8
M- 11111 1	6
N- 11111 11111 1111	14
O- 11111 11111 11111 11	17
P- 11111	5
Q- 1	1
R- 11111 11111 1111	14
S- 11111 11111 111	13
T- 11111 11111 111	13
U- 11111 1	6
V- 111	3
W- 111	3
X- 1	1
Y- 1111	4
Z- -	-
Total.....	200

(b)

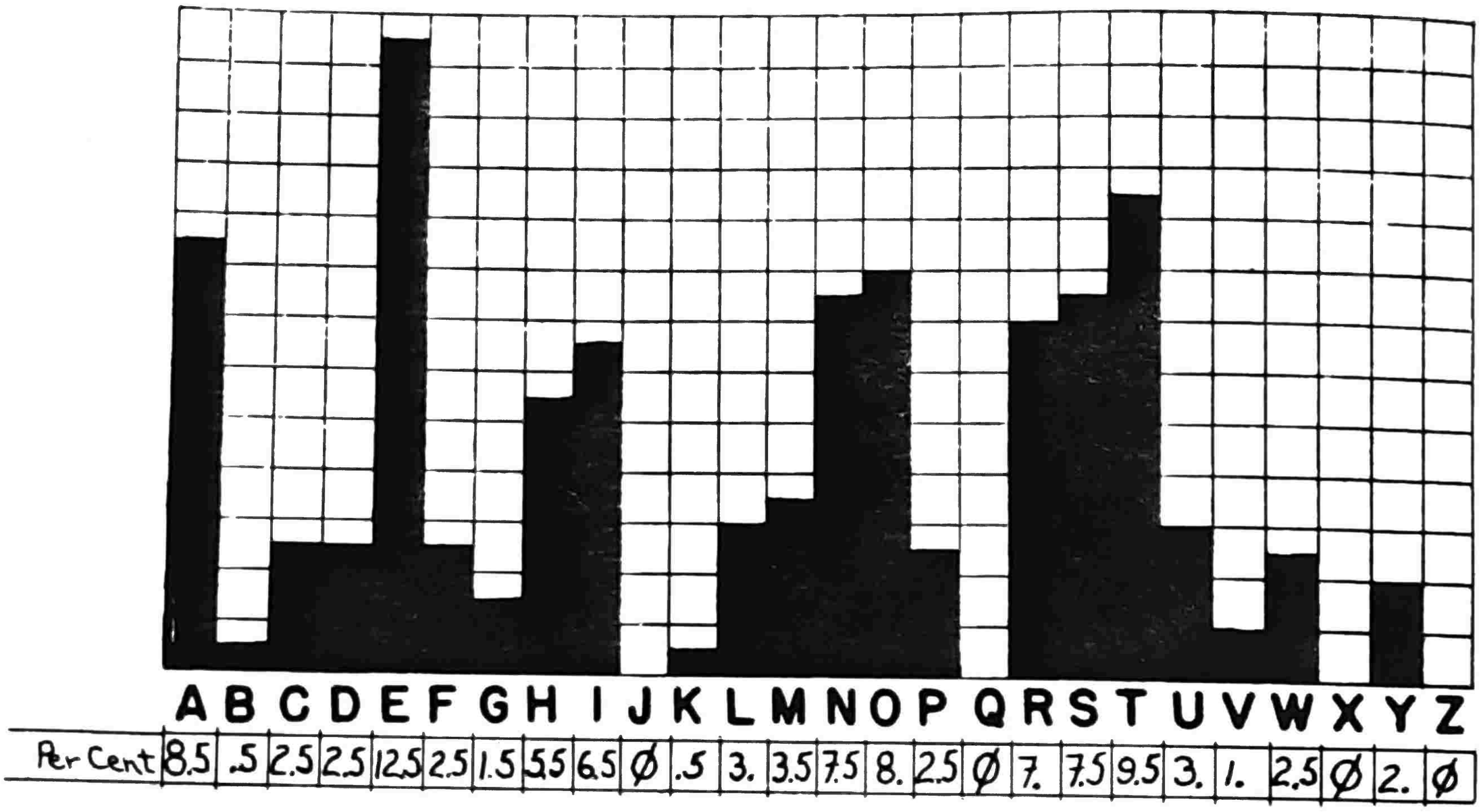
In Order of Frequency	
E- 11111 11111 11111 11111 11111 1	26
O- 11111 11111 11111 11	17
A- 11111 11111 11111 1	16
N- 11111 11111 1111	14
I- 11111 11111 1111	14
R- 11111 11111 1111	14
S- 11111 11111 111	13
T- 11111 11111 111	13
D- 11111 111	8
L- 11111 111	8
H- 11111 111	8
U- 11111 1	6
C- 11111 1	6
M- 11111 1	6
P- 11111	5
Y- 1111	4
F- 1111	4
G- 1111	4
W- 111	3
B- 111	3
V- 111	3
K- 11	2
X- 1	1
J- 1	1
Q- 1	1
Z- -	-
Total.....	200

FIGURE 1

In (a) of Figure 1, note A, E, and I, three vowels having a high frequency of occurrence, are each separated by three consonants of medium or low frequency. Then come four medium and low frequency consonants followed by N occurring quite often and next to N the very high frequency vowel O. Succeeding O are the two low frequency consonants P and Q, and after Q, we find three high frequency consonants R, S, and T. At the end is a blank, representing Z, which rarely occurs. These

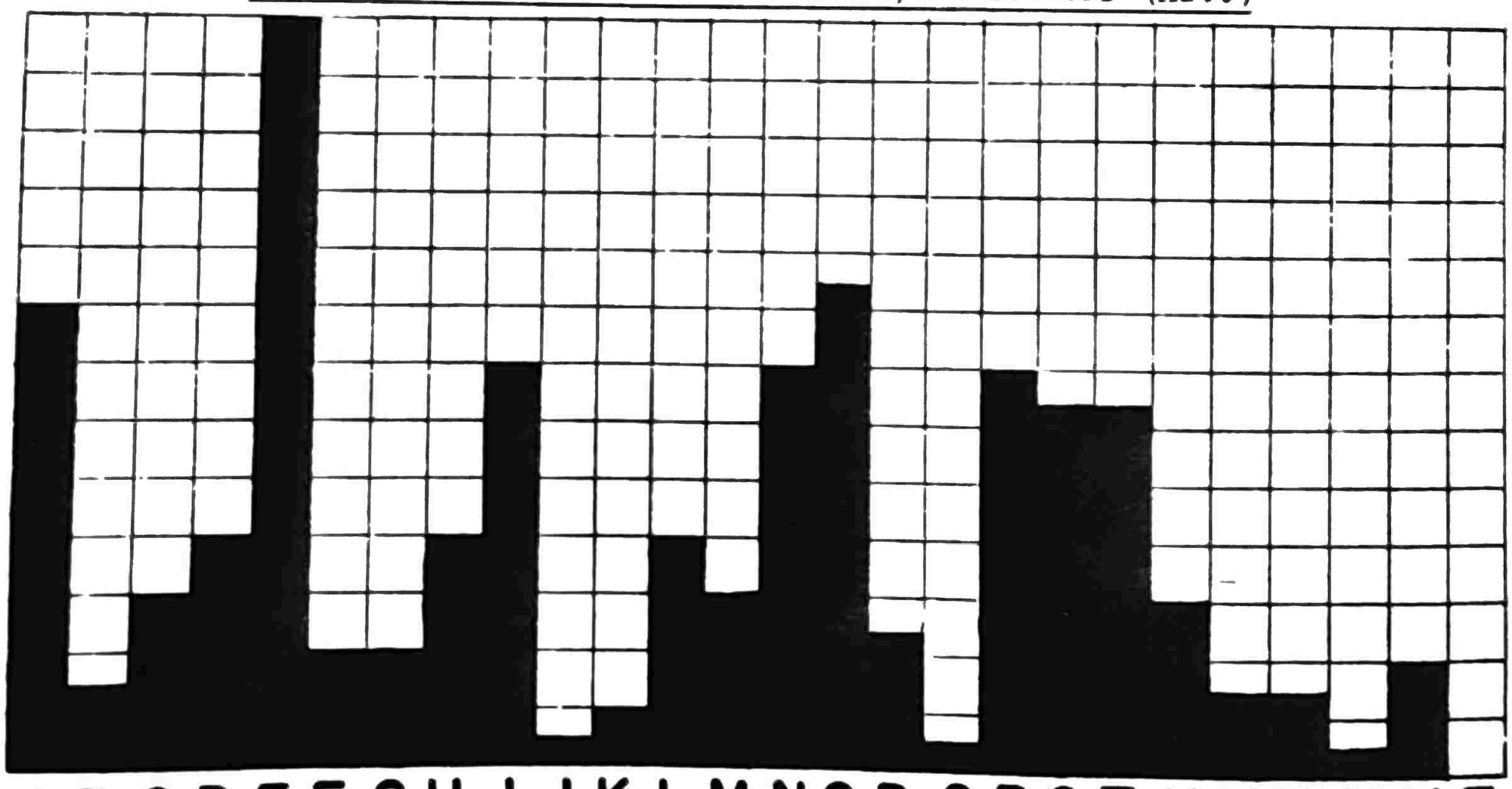
FREQUENCY GRAPHS - ENGLISH

a. Literary Text - Count of 1,030 Letters



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Per Cent	8.5	.5	2.5	2.5	12.5	2.5	1.5	5.5	6.5	Ø	.5	3.	3.5	7.5	8.	2.5	Ø	7.	7.5	9.5	3.	1.	2.5	Ø	2.	Ø

b. Telegraphic Text - Count of 10,000 Letters (Hitt)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Per Cent	8.	1.5	3.	4.	13.	2.	2.	4.	7.	.5	1.	4.	3.	7.	8.5	2.5	.5	7.	6.5	6.6	3.	1.5	1.5	.5	2.	Ø

FIGURE 2

spacial relationships, as graphically shown in Figure 2, are of great importance and their value will be brought out as we proceed. Note the "crests" showing the high frequency letters and the "troughs" representing those of low frequency. The charts may be compared with similar ones prepared for any given cryptogram. Such comparisons will often be of great assistance in determining whether the cipher is transposition or substitution and if substitution whether monoalphabetic or polyalphabetic, besides helping to identify individual letters of the cipher alphabet.

In English telegraphic text the normal percentages are as follows:

Vowels A, E, I, O, U	40.08%.
Consonants L, N, R, S, T	30.77%.
Consonants J, K, Q, X, Z	2.25%.
Remaining 11 letters, B, C, D, F, G, H, M, P, V, W, Y	26.90%.

The usual practice is to take approximate figures: Vowels - 40%; L, N, R, S, T - 30%; J, K, Q, X, Z - 2%; and the remaining 11 letters - 28%. The letter E usually may be counted on to appear 13 times out of every 100 letters. Variations from the normal percentages within 5% either above or below are not uncommon. IT MUST BE EMPHASIZED THAT CHARTS, GRAPHS, AND TABLES SHOWING FREQUENCY DISTRIBUTION ARE TO BE CONSIDERED AS GUIDES. Each cryptogram will have its own peculiarities and frequency data alone are not necessarily conclusive.

Tables of single letter (uni-literal) frequency must be supplemented by information regarding digraphs and trigraphs, data on initial and final letters of words, etc.

Figure 3 shows a table of digraphs for English literary text. The first letter of the pair is found in the columns of letters at the left and right, while the second letter of the pair is given in the rows at the top and bottom of the table.

DIGRAPHIC FREQUENCY CHART - ENGLISH LITERARY TEXT

Second Letter

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A		1	4	3		4			3		1	3	1	23		2		15	12	5	4	2			3	
B					2										1											
C	3				5			2	2		1				9		1			4						
D	3				7	2			2			1			1	1			3	3	1				1	2
E	8		3	2	6	2	1	5				4	6	15	4	5		28	4	9	1	2	8		3	3
F	1		1		2	2		2	2						4			2	1	4	1	1			3	
G	2				2			6				1		1						3	1					
H	5		1		28	1		1	9						3				1	1			3			
I	1		3	1	2	1	3					5	3	18	1			3	15	7		3				
J																										
K	3				1									2												
L	4			4	7	2			4			5			1					1	1		1			
M	12				11				3				2		1	3			2	1	1		1			
N	2			13	4	1	9	1	4				2	3	9				4	11		1	2		3	
O	1		1			8					1	1	1Ø	6	2	8		5	3	7	18	5	1		2	
P	2				5			1	1			4			3			6	1	1	2		1			
Q																					1					
R	7		2	3	13	2		2	7		2		4		6	1		1	7	6	1		4		3	
S	9	1	3		12				7				3	2	8	2		3	7	15			2			
T	9	1			1Ø	1		29	7			2			11			3	1Ø	1Ø					4	
U	1	1		1			3	1				3	4	4		2		5	2	5			1			
V					12			3																		
W	1Ø							3	6				1		4			1		1						
X	1																									
Y	2					1						1	1	1	11				2	2			1			
Z											1					1							1			

Second Letter

FIGURE 3

This compilation is based on a count of 1,030 letters, taken from "False Dawn," in "Plain Tales from the Hills," by Rudyard Kipling.

Let us try out the following message and see what the three types of frequency tables will show.

"WHETHER ON ACTIVE SERVICE OR MERELY AT DRILL THE COMMANDER OF AN ADVANCE GUARD OUTPOST OR REAR GUARD MUST EXERCISE JUDGMENT AND MAKE HIS DISPOSITIONS IN ACCORDANCE WITH THE NATURE OF THE GROUND AND THE REAL OR SUPPOSED CIRCUMSTANCES OF WARFARE UNDER WHICH HE IS ACTING."

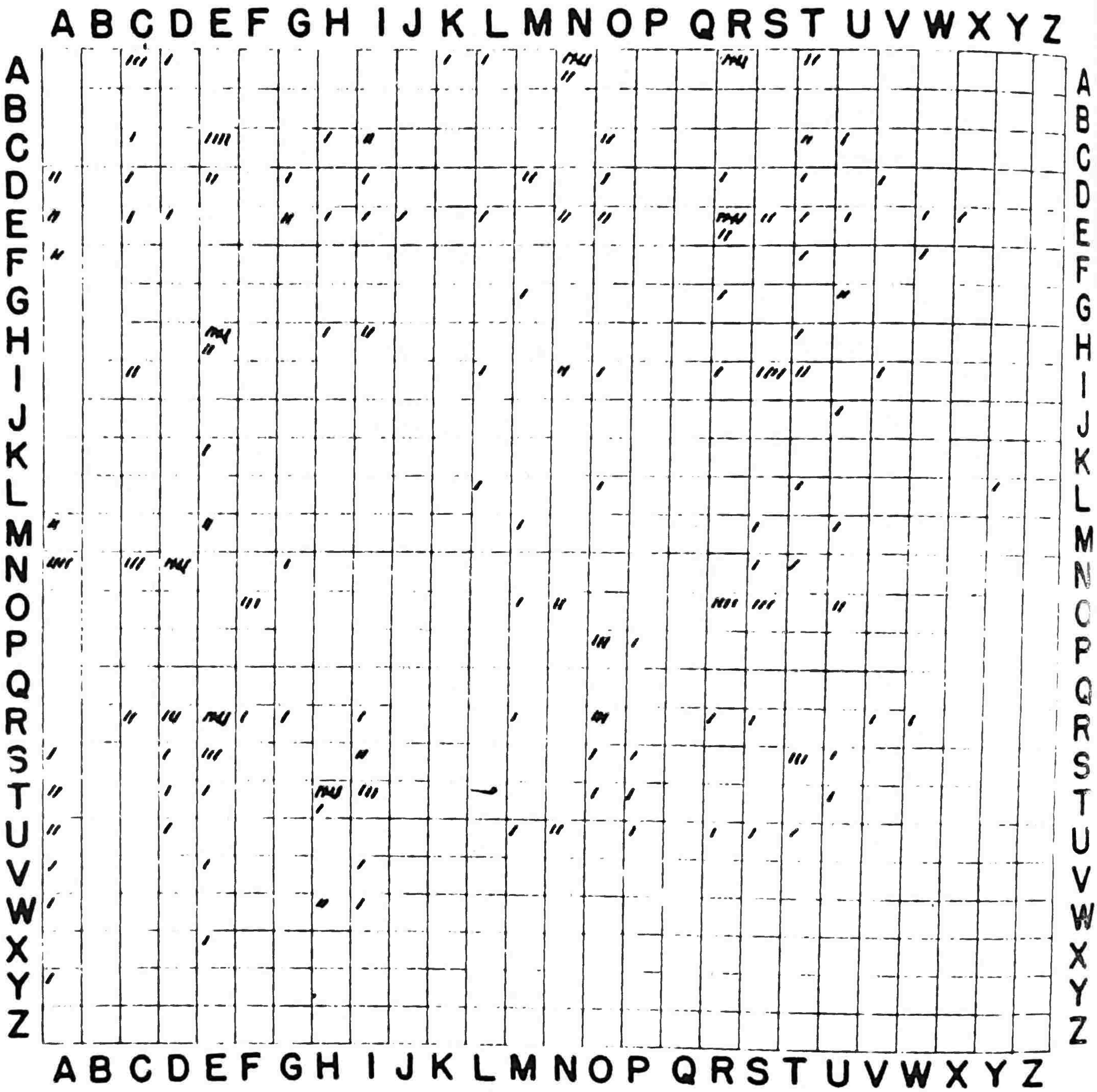
The sentence contains 223 letters and we can compare the single letter frequency table with that of normal literary text for 200 letters.

Normal Text - 200 Letters		Military Text - 223 Letters	
A	16	A	20
B	3	B	0
C	6	C	13
D	8	D	13
E	26	E	27
F	5	F	4
G	4	G	5
H	11	H	11
I	14	I	14
J	0	J	1
K	1	K	1
L	8	L	4
M	5	M	7
N	14	N	15
O	16	O	15
P	4	P	4
Q	0	Q	0
R	12	R	21
S	13	S	13
T	18	T	16
U	6	U	10
V	2	V	3
W	4	W	4
X	0	X	0
Y	4	Y	1
Z	0	Z	0
	200		223

The "crests" and "troughs" in both tables are very similar, although there are some variations in the number of times some of the letters occur. For all practical purposes the military sentence may be considered as following closely the literary frequency. Note the particular variations in the letters B, C, D, L, R, T, U, and Y, which, in the military text, would change the order in which these letters occur. A digraphic chart on page 20 and a trigraphic table on page 21 reveal additional information.

DIGRAPHIC CHART

Second Letter



First Letter

FIGURE 4

List of Digraphs Occurring More Than Once:

AN-7,	ER-7,	HE-7,	TH-6,	AR-5,	ND-5,	RE-5,	CE-4,	IS-4,	NA-4,
OR-4,	AC-3,	NC-3,	OF-3,	OS-3,	PO-3,	RD-3,	RO-3,	SE-3,	ST-3,
TI-3,	AT-2,	CI-2,	CT-2,	DA-2,	DE-2,	DM-2,	EA-2,	EC-2,	EN-2,
ED-2,	ES-2,	FA-2,	GU-2,	HI-2,	IC-2,	IN-2,	IT-2,	MA-2,	ME-2,
ON-2,	OU-2,	RC-2,	SI-2,	TA-2,	UA-2,	UN-2,	WH-2.		

TRIGRAPHIC FREQUENCY TABLE No._____

A	N Y M F N V U E U T M N D N D E T W F S C T N N D N R R R N K C N T N L N R R C	2Ø
B		Ø
C	A I E N R A C N D R N I A T E O E I C O E I U E H T	13
D	T N A R R U N S R N N E N R E V O M G M I A A T C E	13
E	H H V S C M R H D C R T X S M K C H R H H R S C R D H T R S R O R L C R G A X R J N H W N O G R A D S U R I	27
F	O O O R A T W A	4
G	E R D E N U U M R -	5
H	W T T E T T T T W C H E E E I T E E E I H E	11
I	T V R C H D S T S W C H E T V C L S S S T O N T R C S N	14
J	E U	1
K	A E	1
L	E I L A Y L T O	4
M	R O M D G D U E M A U E A S	7
N	O A A A M A O I A E U A A U I A D A C N D S A C A D D C D G	15
O	R E C R D F T F I C E R L P S N R M F U S R S N R F U R S F	15
P	T S U P O O P O	4
Q		Ø
R	E E O E D E A O R A A E O U G E O I A A E O V M E I O D R E G D C D E O E S C F E W	21
S	E O U I I I O N R O M E I E T T E D F I I U E T O A	13
T	E C A L U S S N I I H A F D S C H I D H F O E A I H H U H H A I	16
U	G O G M J T O S C E A T A S D R N P M N	1Ø
V	I R D E I A	3
W	- E F R H I A H	4
X	E E	1
Y	L A	1
Z		Ø
	Total	223

REMARKS:

FIGURE 5

Order of Frequency:

E R A T O N I S H D C U M G L F P W V Y X J K B Q Z.

Trigraphs occurring more than once are:

THE-5, ANC-3, AND-3, NCE-3, POS-3, ACT-2, ARD-2, CTI-2, DAN-2, DER-2, ERE-2, ERO-2, GUA-2, HER-2, NAC-2, NCE-2, NDE-2, REA-2, TAN-2, UAR-2, and UND-2.

Note that the first letter in the example is W. Upon its first occurrence there is no letter preceding it. This is shown on the trigraphic chart by placing a dash (-) above the line, thus: $W\overline{-}--$. The last letter in the example is G and the fact that no letter follows it is indicated by placing a dash below the line: $G--------\underline{N}$. If x is the number of single letters in the message, (in this case - 223), the total number of real trigraphs will be x - 2, (in the example it is 223 minus 2, or 221).

An alternative form of a trigraphic chart is given below.

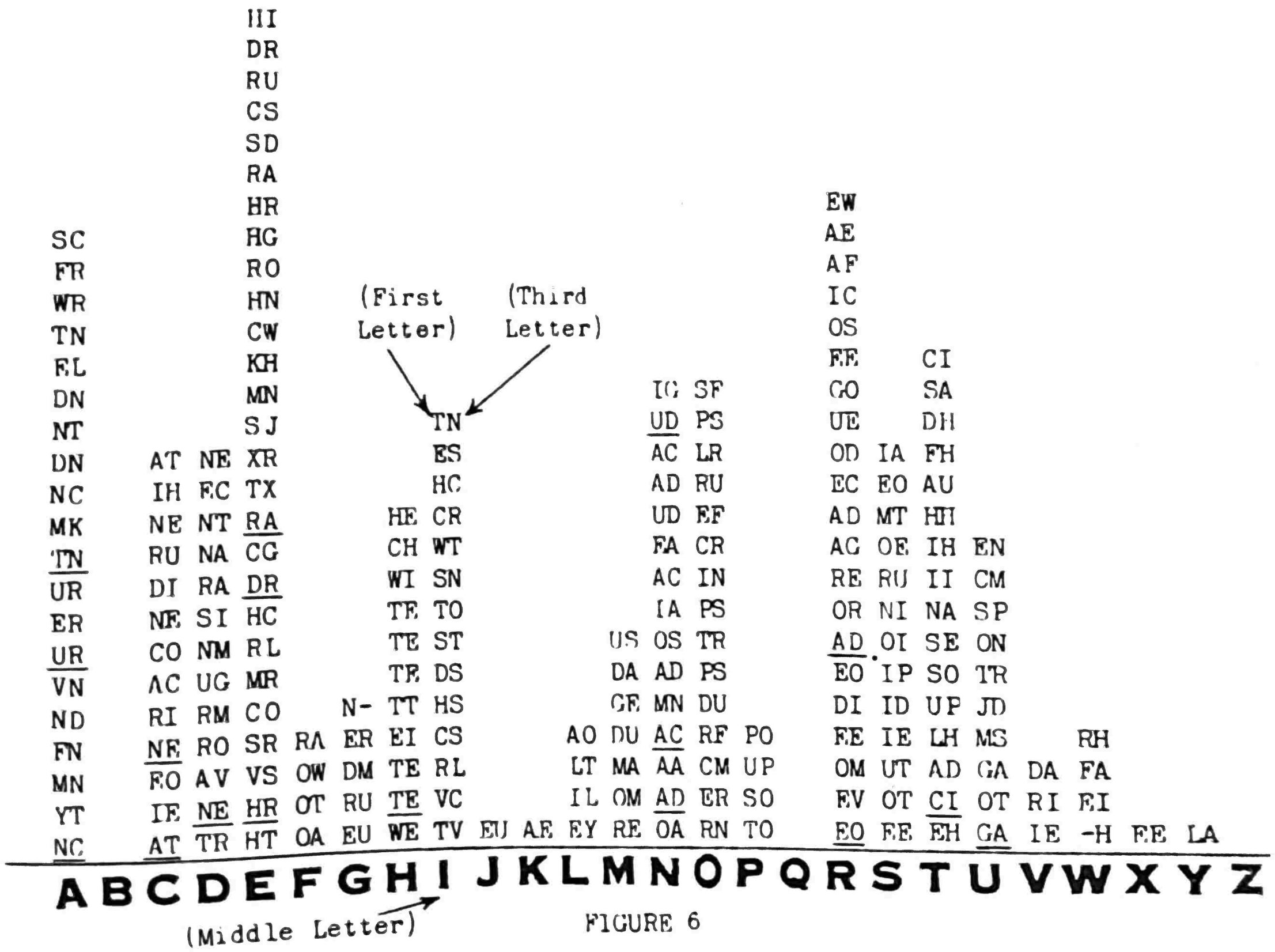

III
DR
RU
CS
SD
RA
HR
SC HG EW
FR RO AE
WR HN AF
TN CW IC
EL KH OS
DN MN FE CI
NT SJ GO SA
DN AT NE XR UE DH
NC IH EC TX OD IA FH
MK NE NT RA EC EO AU
TN RU NA CG AD MT HI
UR DI RA DR AC OE IH EN
ER NE SI HC RE RU II CM
UR CO NM RL OR NI NA SP
VN AC UG MR AD OI SE ON
ND RI RM CO EO IP SO TR
FN NE RO SR DI ID UP JD
MN EO AV VS FE IE LH MS RH
YT IE NE HR OM UT AD GA DA FA
NC AT TR HT EV OT CI OT RI EI
EO EE EH GA IE -H EE LA
(First Letter)
(Third Letter)
TN
ES
HC
HE CR
CH WT
WI SN
TE TO
TE ST
TE DS
TT HS
N-
ER EI CS
RA
OW DM TE RL
OT RU TE VC
EU AE
IG SF
UD PS
AC LR
AD RU
UD EF
FA CR
AC IN
IA PS
US OS TR
DA AD PS
GE MN DU
AO DU AC RF PO
LT MA AA CM UP
IL OM AD ER SO
EY RE OA RN TO
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
(Middle Letter)

FIGURE 6

Appearances of single letters, digraphs, trigraphs and words may be termed "Frequency distribution." This distribution varies with each individual text under consideration.

Much of the labor in cryptanalysis is in the preparation of frequency data for each cryptogram. Constant practice is required to develop accuracy and speed.

Chapter IV

RULES FOR THE USE OF CODES AND CIPHERS

General Marcel Givierge, one of the Chiefs of the French Code and Cipher Section during World War Number One, stated: "Cryptograph well or do not cryptograph at all." A cryptogram that has been poorly or carelessly enciphered or encoded will give the enemy considerable help in breaking it. Much of the success of cryptanalysts is achieved because of the errors or carelessness of their opponents.

The following rules should be rigidly observed:

1. Study carefully, understand thoroughly, and adhere strictly to, the instructions in each code book or supplied with each cipher.

2. Exercise care to prevent the loss or compromise of a code book or a cipher key. If a code book or cipher key is lost or compromised, report the fact promptly to your superior.

3. Messages to be sent in code or cipher should be as short as possible. Sometimes it is practicable to break a long message up into short ones to be transmitted separately. Long messages aid the enemy in their solution.

4. Never repeat a message in a code or a cipher other than the one in which it was originally transmitted. It is much better to paraphrase the original message and send it again in the original code or cipher used. In paraphrasing, do not expand the text. Re-arrange the sentences or paragraphs. Change the order of the subject, predicate and modifying phrases or clauses. Alter the wording by using synonyms.

5. Never send in code or cipher a message which has been transmitted previously in clear.

6. Never send in clear a message previously sent in code or in cipher.

7. Never mix clear text and secret text. If a message is to be sent in code or cipher, the entire message must be encoded or enciphered. Spell out numbers and signs of punctuation.

8. To avoid errors, capital letters should be used in writing cryptograms. Cipher messages are usually transmitted in groups of five letters. The groups of the code text correspond to the code groups shown in the code book. Code messages are sometimes transmitted in figure groups, usually three, four, or more often, five figures in a group. In writing figures use the Arabic numerals and be careful to distinguish between 1 and 7 and between 3 and 5.

9. In cryptographing and decryptographing all work must be checked by another person for accuracy. Of course, this can only be done if sufficient time and personnel are available.

Hints and Suggestions

1. UNDERSTAND and ANALYZE each problem carefully before you start to work on it.
2. Check each step as you go along.
3. If you get a "hunch," make a note of it, but do not act upon it before your analysis is completed unless you are certain that it is logical and not offset by other factors.
4. Cryptanalysis is by no means a "one-man" job. Your work will have to be read by others. Write legibly. Print all letters in Roman Capitals and all figures in Arabic Numerals.
5. Do not crowd your work. Use sufficient space, but do not waste paper.
6. Indicate all errors. If a letter has been wrongly enciphered or deciphered, it may be underlined - perhaps with a different-colored pencil.
7. Printers have a term - "Slugging." Each job is assigned a number and this number accompanies the job until the latter is completed. Each problem, its work-sheets, and the solution, should bear the same number. If sheets become separated, or are mislaid, they can be identified and put back in their proper places quickly.

8. Make a frequency chart, preferably trigraphic, for each cryptogram and keep it for comparison with those for other cryptograms that may be received later. This may save much time and labor.

9. Whenever possible, reconstruct both the enciphering and the deciphering alphabets, key-words or key-phrases and numerical keys, and matrices if these are employed. They may prove very helpful in the solution of other cryptograms.

10. DON'T RUSH! TAKE YOUR TIME, BUT DON'T WASTE IT! HASTE VERY FREQUENTLY CAUSES ERRORS!

Samples of Printing:

ROMAN CAPITALS

ABCDEFGHIJKLMNOPQRSTUVWXYZ

ARABIC NUMERALS

1234567890

Some Mechanical Aids

1. A set of ANAGRAM LETTERS, printed on squares of cardboard or wood and usually available at the five-and-ten-cent stores, will be found quite useful in anagramming to find a key-word and to solve certain types of ciphers.

2. A set of POLYALPHABET STRIPS. These are narrow strips of cardboard or wood, one-quarter to five-sixteenths inches wide, having the letters of the alphabet printed twice in normal sequence from top to bottom of each strip. They can be slid against one another and are especially useful in "running down the alphabet." 10 are needed, but 20 are better. Two or more, bearing double alphabets in reversed order, are also desirable. Each strip must be long enough to take 52 letters and have margins of at

least one-half to three-quarters of an inch at both top and bottom. The strips may be cut from heavy Bristol Board or cardboard, but the cutting should be done with a razor blade or a very sharp pen-knife, otherwise the strips may curl and be difficult to handle.

3. A slide-rule, bearing the normal and reversed alphabets in a double sequence on both the fixed and sliding portions, will be found extremely useful.

4. Ruler, 12 to 15 inches long.

5. A transparent, right-angled, triangle with equal legs, such as draftsmen use, can be employed to advantage when dealing with square tables such as the Vigenere.

6. Additional supplies are essential, particularly:

    a. Quadrille-ruled, ("cross-section" or "graph"), paper having 4 or 5 squares to the inch. 4 squares to the inch is preferable as there will be no need for crowding or for printing letters or figures so mall that they will be difficult to read.

    b. Black and colored, (red, blue, green, etc.), pencils.

    c. Erasers.

    d. Pair of scissors.

    e. (Optional). Alphabet Rubber Stamp, to fit cross-section paper used, while desirable is not actually essential.

    f. (Optional). A blank slide-rule and rubber cement for temporarily mounting alphabets.

Chapter V

ON THE ANALYSIS OF CRYPTOGRAMS

When a cryptogram is received for solution, much valuable time will often be saved if the preliminary steps in analyzing the problem are taken in a careful and orderly manner. The method of analysis may be compared roughly to that of making an "estimate of the situation" before coming to a decision as to the action to be taken. Whether in civil or military affairs, making a decision to do something is always preceded by a summing-up of the information at hand and the possible courses of action available. In many cases the summing-up and the decision are so quickly made that they appear to be accomplished simultaneously or the decision seems to be almost instantaneous. Although many instances will arise where a detailed written analysis may seem to be superfluous, the practice of following a definite routine is an excellent one. The requirement of a regular and rather fixed procedure, as in the case of drawing up a military field order, develops a habit of thinking that helps to eliminate the chance of overlooking some of the essential factors. Every scrap of information from all possible sources concerning the cryptogram and the circumstances under which it was sent and obtained should be utilized.

Points to be considered in the preliminary analysis of a cryptogram comprise:

1. Map Reference, if any. This may include information about the particular area, or a city and its environs, in which the cryptogram originated or where the action is taking place.
2. Collateral information concerning the enemy and our own forces. This may comprise all that is known about the sender and the addressee of the cryptogram, his organizations, and his past and present activities in each case. It should also include data on our own organizations and activities. All of this is valuable in determining or attempting to discover the probable subject matter of the cryptogram and may be helpful in securing the

"entering wedge" that will lead to a solution.

3. Information about the message itself.

   a. The number of the cryptogram, (if any).

   b. The date and the hour that it was sent on its way.

   c. The name or other designation of the addressee and the place of destination.

   d. The name or designation of the sender and the point of origin.

   e. How, when, and by whom, the cryptogram was obtained.

   f. What means of transmission was being used when it was obtained.

   g. The probable language of the plain text of the cryptogram.

   h. The character of the cryptographic text, (letters, figures, symbols, or other forms).

   i. The character of the groups in the cryptographic text, (5-letter, 3-figure, etc.).

In Figure 7, on Page 29, is shown one form of a Cryptanalysis Data Sheet that has been devised for the recording of the information about cryptograms. During the preliminary analysis all available data should be entered on this form, down to and including Item C-12 (Kind of Groups). The remaining items, 13 to 18 inclusive, should be filled in as the analysis progresses and final solution is obtained. Even if it seems temporarily impossible to complete the solving of the message, all records should be preserved, because at some future time another cryptogram of similar type may be solved and thus lead to the "breaking" of the uncompleted problem. All data sheets should be classified and filed for future reference and for making comparisons with those on other cryptograms that may be received at a later time.

CRYPTANALYSIS DATA SHEET Number:_____

A. MAP REFERENCE: Date:_____

B. COLLATERAL INFORMATION:

1. Enemy Forces:

2. Own Forces:

C. CRYPTANALYTICAL and MESSAGE DATA:

1. - Message No._____ 2. - Date_____ 3. - Hour_____
4. - Addressee_____
5. - Sender_____
6. - Obtained by_____
At_____ Date_____ Hour_____
7. - How being sent when obtained (a) Radio_____(b) Telegraph_____
(c) Telephone_____(d) Buzzer_____(e) Helio_____(f) Lantern_____
(g) Flags_____(h) Panels_____(j) Pigeon_____(k) Motorcyclist_____
(l) Cyclist_____(m) Foot Messenger_____(n) Mounted Messenger_____
(o) Mail_____(p) Miscellaneous_____
8. - Probable Language_____
9. - Character of Text (a) Letters_____(b) Numerals_____(c) Symbols_____
(d) Miscellaneous_____
10. - No. of characters in text_____ 11. - No. of Groups_____
12. - Kind of Groups_____
13. - Percentages E_____, AEIOU_____, LNRST_____, JKQXZ_____, Others_____
14. - Probable class of message (a) Code_____, (b) Cipher_____, (c)_____
15. - Probable type of Code (a)_____-Letter (b)_____-Numeral_____
(c) Enciphered_____(d) Miscellaneous_____
16. - Probable class of Cipher (a) Transposition_____(b) Substitution_____
17. - Probable Type of Transposition (a) Columnar_____(b) Route_____
(c) Complete figure_____(d) Incomplete Figure_____(e) Grille_____
(f) Single_____(g) Double_____(h) Letters_____(j) Words_____
(k) Miscellaneous_____
18. - Probable type of Substitution (a) Monoalphabetic_____(b) Polyalphabetic_____(c) Monographic_____(d) Polygraphic_____
(e) Standard Alphabets (e-1) Direct (e-2) Reversed
(f) Mixed Alphabets (f-1) Keyword Mixed (f-2) Random
(g) Miscellaneous

D. REMARKS:

FIGURE 7

Methods of Analysis

Frequential Analysis Versus the Probable Word

There are two schools of thought among cryptanalysts. One of these relies upon the analytical method based upon frequency data. The other places its main reliance upon the intuitive method founded upon a "probable word" in the plain text of the cryptogram.

Let us take up first the case for the "probable word." Every cryptogram presupposes both a sender and an addressee. Knowledge of who the addressee and the sender are and information concerning their activities will generally be sufficient for the cryptanalyst to form an idea as to the type of subject matter dealt with in the cryptogram. A military communication will not treat of the same matters as a message exchanged between business firms, or social correspondence, or a letter from one criminal to his pal. Military messages are likely to contain such words as ATTACK, ARTILLERY, INFANTRY, DIVISION, and REAR. In all of these five words and in a great many others there are letters that are repeated. ATTACK and ARTILLERY contain the doublets TT and LL. In the word INFANTRY the two N's are separated by two letters, while in the word DIVISION there are three I's with single letters between them. REAR has an R at the beginning and as a final. All of these are "pattern words." The technical term is "idiomorphs." In a literary text such words as RECEIVE, TOMORROW, and UNUSUAL occur, also AGAIN, BEGINNING, and MONOTONOUS, as well as INDEED, NOON, etc. A cryptogram may contain repeated groups or repeated letters that have appearances similar to the "pattern words" that appear in plain text. By means of the trial and error method the cryptanalyst endeavors to fit his supposed word to the pattern that may represent it in the cipher text. If he fails with one "probable word," he can try others. This method is especially effective when a cryptogram is so short that analysis by frequency would not yield a sufficient distribution to serve for identification of letters according to the usual frequencies. Sometimes the identification of the cipher equivalents of a

pattern word by the "probable" method enables the cryptanalyst to gain an entering wedge and solve the problem perhaps more rapidly than he would have with the frequency analysis method. This system of attack is most effective when the cipher is of the single substitution or monoalphabetic type. Intuition, or, as we sometimes call it, "intelligent guess-work," is a prime factor in the probable word method.

The idea of a "normal" frequency is not a mathematical exactness. Nevertheless, the characteristics of the language concerned are of definite and practical value in the use of the frequency analysis method. The formation of the plurals, endings of verbs in the various tenses, the comparison of adjectives, and the other applications of the rules of grammar, all contribute to the identification of single letters, digraphs and trigraphs. The articles, conjunctions, prepositions, and other words that are usually omitted in telegraphic communications form a class distinctive in themselves and often easily recognized or isolated. Indeed, every peculiarity of the particular language involved or suspected in a cryptogram yields a clue to the cryptanalyst. Prefixes and suffixes serve to differentiate certain letters or groups of letters from others. Preparation of frequency charts graphs and tables is a long, arduous and exacting process, but it is an indispensable one for the cryptanalyst. Solution of a cryptogram by analyzing the frequencies requires more time than the probable word method, but it is more exact and more accurate.

The best results are undoubtedly attained by the use of both methods in conjunction. In fact they supplement one another remarkably well.

Chapter VI

THE TWO PRINCIPAL CLASSES OF CIPHERS

DISTINGUISHING BETWEEN TRANSPOSITION AND SUBSTITUTION

The two principal classes of ciphers are transposition and substitution. In a transposition system the original letters or words of the plain text message are simply rearranged in a different order, while in a substitution the original elements of the text are replaced by different ones, which may be letters, figures, or symbols. Frequency charts for a cryptogram enciphered by transposition will closely approximate those for normal text of the particular type involved - literary, newspaper, telegraphic, military, etc. Frequencies for a message, the secret text of which is produced by substitution, will differ radically from the normal. Each class of cipher has its own distinguishing characteristics.

If a cryptogram consists of symbols or figures, or a combination of both, either with or without letters, the cipher is obviously of the substitution class. If the cryptogram is composed of letters only, but there is a frequent occurrence of the rare consonants J, K, Q, X, and Z, the system of encipherment is also clearly substitution.

When the cryptogram consists of letters only and the frequencies for the five vowels, A, E, I, O, and U, for the often recurring consonants, L, N, R, S, and T, and for the rarely encountered consonants, J, K, Q, X, and Z, closely resemble those for the normal text, and in addition the percentages for the individual letters are approximately normal, then we may assume that a transposition system has been used. The normal percentages of 40% for the five vowels, 30% for L, N, R, S, and T, and 2% for the rare consonants, with 28% for the remaining 11 letters, should not vary more than 5% either way. These figures are, then, 35% to 45% for the vowels, 25% to 35% for L, N, R, S, and T, with zero to about 7% (which is unusual) for the rare consonants, and 23% to 33% for the other 11 letters. The indi-

vidual frequencies of letters must also be checked against the normal, because sometimes in a substitution cipher the cipher alphabet has been so prepared that vowels are substituted for vowels and the high-frequency and rare consonants, as well as the remaining letters, have simply been replaced by others within their respective groups. Therefore, merely checking the data for the groups of vowels and consonants, without paying careful attention to the percentages for the individual letters, will result in sometimes drawing false conclusions and may cause considerable waste of time and effort.

Let us take two cryptograms having the same plain text, totalling 215 letters and determine whether transposition or substitution was the means of encipherment in either case.

Example 1.

UNERA DYLAY RTNOV ONPEA LPETD REMDL FRIEL EETRC
FAFTD RYRRH YARNO IESIS SETPO ERCNO NCTRH TMSTL
TNTAA NIUPU TEINE LNCNR AGOSF ELLDO GERGR AILLA
SLMDT EESCB IHOUO ALSCP AUGAL MVVYT OOTSE NDNAI
BSVTI YNSIT BAUPE EOOCT EOLTR ERNNN EWROE ONEAT
SPMOS KFRED LIMUA.

First glance over the cryptogram tells that there is only one appearance of any of the five rare consonants, K in the next to the last 5-letter group. A further quick of the first line of 40 letters gives the following: Vowels, A,E,I,O,U - 14 or 35%; Consonants L,N.R,S,T - 15 - 37.5%; Consonants J,K,Q,X,Z - Ø; remaining 11 letters - 11 or 27.5%. These two facts point to transposition. The percentages for the entire 215 letters are:

	CRYPTOGRAM	NORMAL	VARIATION
Vowels, A,E,I,O,U, ...	35.9%	4Ø.Ø%	Minus 4.1%
Consonants, L,N,R,S,T, ...	38.7%	3Ø.Ø%	Minus 8.7%
Consonants, J,K,Q,X,Z, ...	Ø.5%	2.Ø%	Minus 1.5%
Letters - B,C,D,F,G,H,M,P,V,W,Y, .	24.9%	28.Ø%	Minus 3.1%
Totals	1ØØ.Ø%	1ØØ.Ø%	----------

While the percentages of the high-frequency consonants are well above the 5% variation, the percentages for the remaining three groups are within the limit of 5% above or below. Peculiarities of the text may be the cause. We care fairly safe in assuming that transposition has been used, but in order to be sure we will check the individual frequencies of the letters in the first three groups.

		CRYPTOGRAM	TELEGRAPH NORMAL	VARIATION
Vowels -	A	8.Ø%	8.Ø%	none
	E	11.7%	13.Ø%	Minus 1.3%
	I	5.1%	7.Ø%	Minus 1.9%
	O	8.Ø%	8.5%	Minus Ø.5%
	U	3.1%	3.Ø%	Plus Ø.1%
Consonants -	L	7.Ø%	4.Ø%	Plus 3.Ø%
	N	8.4%	7.Ø%	Plus 1.4%
	R	8.4%	7.Ø%	Plus 1.4%
	S	6.Ø%	6.5%	Minus Ø.5%
	T	8.9%	6.5%	Plus 2.4%
Consonants -	J	zero	Ø.5%	Minus Ø.5%
	K	Ø.5%	1.Ø%	Minus Ø.5%
	Q	zero	Ø.5%	Minus Ø.5%
	X	zero	Ø.5%	Minus Ø.5%
	Z	zero	zero	none

Considering the shortness of the text these variations are not too great to preclude the decision that a transposition system has been employed. As we have mentioned at the beginning of Chapter III, some variations from the normal are to be expected.

Example 2.

MBOSH JKEBO JCGOA SHJBC HKLSH BPCGM ABKCC NSQSH
YKLAB UACQE XGSTH CAPED KMSHO EQAEJ CHEBO THCAD
EHKGY XMHBS ODGEB SJMUU SJKJD SSOYS TTCHK KCHSP
CKSHE GGEHK APGSJ EJDCK SBKAE GAOSB KATAP EKACB
JAYDS HJCBB SGXMJ YCBDM HSGYA BKSGG AUSBP SAEKK
SHJPE GGELE BJKRC.

A quick survey discloses the presence of the five rare consonants, J, K, Q, X, and Z, in considerable numbers. Checking the first 4Ø letters comprising the first line yields:

		CRYPTOGRAM	TELEGRAPH NORMAL	VARIATION
Vowels, A, E, and O, . . 7 (I and U missing)		17.5%	29.5%	Minus 12.Ø%
Consonants, L, N, and S, . . 6 (R and T missing)		15.Ø%	17.5%	Minus 2.5%
Consonants, J, K, and Q, . . 7 (X and Z missing)		17.5%	2.Ø%	Plus 15.Ø%
Consonants, B,C,G,H,M, P, (D,F,V,W,Y, are missing)	20	5Ø.Ø%	15.Ø%	Plus 35.Ø%

These figures point definitely to a substitution system. Figures for the entire cryptogram of 215 letters give us the following results.

	CRYPTOGRAM	NORMAL	VARIATION
Vowels - A,E,I,O.U, ...	22.3%	4Ø.Ø%	17.7% Minus
Consonants - L,N,R,S,T, ...	16.3%	3Ø.Ø%	13.7% Minus
Consonants - J,K,Q,X,Z, ...	18.1%	2.Ø%	16.1% Plus
Consonants - B,C,D,F,G,H,M,P,V,W,Y	43.3%	28.Ø%	15.3% Plus

There is no need to check further. Our cryptogram is clearly of the substitution type.

This chapter has dealt with the distinguishing of the two main classes of cipher, transposition and substitution. As we proceed further along our way we will take up the problems of recognizing the various types belonging to each of these principal classifications.

Chapter VII

THE TRANSPOSITION SYSTEMS

Route Transposition

Transposition ciphers may be compared to "jig-saw" puzzles. All of the original elements of the plain text, (letters or words), are present in their usual forms but the order in which they are arranged has undergone a change. Usually the individual letters of the plain language message are re-arranged according to a definite system, route or pattern. Occasionally pairs or other sets of letters, and, more rarely - entire words, form the elements for re-arrangement.

Nearly all transposition ciphers may be classified as belonging to what is known as the "route" type. In this category the systems of re-arrangement follow specific routes.

Many different geometric designs may be employed, such as - rectangles, triangles, pentagons, and trapezoids, but the most popular pattern is the rectangle or parellelogram.

Two processes, or "steps," are involved: inscription and transcription. The message may be inscribed within the figure according to the usual manner of writing - from left to right - or it may be written in the form of any one of the routes that the correspondents may agree upon. Transcription of the letters to form the secret text is made to follow the specific route agreed to for each message or series of messages.

With the plain text message, WILL MEET YOU TONIGHT AT NINE, below are three examples.

(1)

Inscription	Transcription
WILLME	WILLME
ETYOUT	ETYOUT
ONIGHT	ONIGHT
ATNINE	ATNINE

Cryptogram

WEOA ITNT LYIN LOGI
MUHN ETTE.

(2)

Inscription	Transcription
WILLME	WILLME
TUOYTE	TUOYTE
ONIGHT	ONIGHT
ENINTA	ENINTA

Cryptogram

WTOE NNUI LOII NGYL
MTHT ATEE.

(3)

Inscription	Transcription
WILEUI	WILEUI
LMTTGA	LMTTGA
EYOHTI	EYOHTI
ONTNNE	ONTNNE

Cryptogram

IUELIW AGTTML ITHOYE
ENNTNO.

The following routes, taken from Colonel Parker Hitt's "Manual for the Solution of Military Ciphers," (1916 and 1918), are standard. The order of the letter of the alphabet indicates the route.

(a) Simple Horizontal

1	2	3	4
A B C D E F	F E D C B A	S T U V W X	X W V U T S
G H I J K L	L K J I H G	M N O P Q R	R Q P O N M
M N O P Q R	R Q P O N M	G H I J K L	L K J I H G
S T U V W X	X W V U T S	A B C D E F	F E D C B A

(b) Simple Vertical

1	2	3	4
A E I M Q U	D H L P T X	U Q M I E A	X T P L H D
B F J N R V	C G K O S W	V R N J F B	W S O K G C
C G K O S W	B F J N R V	W S O K G C	V R N J F B
D H L P T X	A E I M Q U	X T P L H D	U Q M I E A

(c) Alternate Horizontal

1	2	3	4
A B C D E F	F E D C B A	X W V U T S	S T U V W X
L K J I H G	G H I J K L	M N O P Q R	R Q P O N M
M N O P Q R	R Q P O N M	L K J I H G	G H I J K L
X W V U T S	S T U V W X	A B C D E F	F E D C B A

(d) Alternate Vertical

1	2	3	4
A H I P Q X	D E L M T U	X Q P I H A	U T M L E D
B G J O R W	C F K N S V	W R O J G B	V S N K F C
C F K N S V	B G J O R W	V S N K F C	W R O J G B
D E L M T U	A H I P Q X	U T M L E D	X Q P I H A

(e) Simple Diagonal

1	2	3	4
A B D G K O	G K O S V X	O K G D B A	X V S O K G
C E H L P S	D H L P T W	S P L H E C	W T P L H D
F I M Q T V	B E I M Q U	V T Q M I F	U Q M I E B
J N R U W X	A C F J N R	X W U R N J	R N J F C A
5	6	7	8
A C F J N R	J N R U W X	R N J F C A	X W U R N J
B E I M Q U	F I M Q T V	U Q M I E B	V T Q M I F
D H L P T W	C E H L P S	W T P L H D	S P L H E C
G K O S V X	A B D G K O	X V S O K G	O K G D B A

(f) Alternate Diagonal

1	2	3	4
A B F G N O	G N O U V X	O N G F B A	X V U O N G
C E H M P U	F H M P T W	U P M H E C	W T P M H F
D I L Q T V	B E I L Q S	V T Q L I D	S Q L I E B
J K R S W X	A C D J K R	X W S R K J	R K J D C A
5	6	7	8
A C D J K R	J K R S W X	R K J D C A	X W S R K J
B E I L Q S	D I L Q T V	S Q L I E B	V T Q L I D
F H M P T W	C E H M P U	W T P M H F	U P M H E C
G N O U V X	A B F G N O	X V U O N G	O N G F B A

(g) Spiral, Clockwise

1	2	3	4
A B C D E F	L M N O P A	I J K L M N	D E F G H I
P Q R S T G	K V W X Q B	H U V W X O	C R S T U J
O X W V U H	J U T S R C	G T S R Q P	B Q X W V K
N M L K J I	I H G F E D	F E D C B A	A P O N M L

(h) Spiral, Counter-clockwise

1	2	3	4
A P O N M L	N M L K J I	I H G F E D	F E D C B A
B Q X W V K	O X W V U H	J U T S R C	G T S R Q P
C R S T U J	P Q R S T G	K V W X Q B	H U V W X O
D E F G H I	A B C D E F	L M N O P A	I J K L M N

The letter A is the starting point in each of these forty rectangles and the letter X shows the end of the route.

The size of the geometric figure employed is, of course, dependent upon the number of letters in the plain text message. In the case of rather long messages it is better to use a comparatively small figure, break the message into two or more approximately equal parts, and encipher and send each part as a separate message.

Let us proceed with a sample message, EVERY OFFICER SHOULD BE FAMILIAR WITH THE PRINCIPLES OF SIGNAL COMMUNICATION AND ELEMENTARY MILITARY CRYPTOGRAPHY. It contains 99 letters. The factors of 99, excluding those in which unity is one, are 3, 3, and 11. These will yield four rectangles: 3 × 33, 33 × 3, 9 × 11, and 11 × 9, the first figure indicating the number of columns and the second figure the number of rows. The two most compact rectangles, or "blocks," are 9 × 11 and 11 × 9. Choosing the simplest method, (a) Simple Horizontal, starting in the upper left corner, we will inscribe the plain text in a rectangle of 9 columns and 11 rows.

E V E R Y O F F I
C E R S H O U L D
B E F A M I L I A
R W I T H T H E P
R I N C I P L E S
O F S I G N A L C
O M M U N I C A T
I O N A N D E L E
M E N T A R Y M I
L I T A R Y C R Y
P T O G R A P H Y

The inscription is complete. Now for the second step, transcription. Using the first case under (b) Simple Vertical, beginning in the upper left, and writing the letters in groups of five for transmission by telegraph, the cryptogram will be:

ECBRR OOIML PVEEW IFMOE ITERF INSMN NTORS ATCIU ATAGY HMHIG

NNARR OOITP NIDRY AFULH LACEY CPFLI EELAL NRHID APSCT EIYY.

With a rectangle 11 × 9, the inscription is:

E V E R Y O F F I C E
R S H O U L D B E F A
M I L I A R W I T H T
H E P R I N C I P L E
S O F S I G N A L C O
M M U N I C A T I O N
A N D E L E M E N T A
R Y M I L I T A R Y C
R Y P T O G R A P H Y

yielding the following cryptogram:

ERMHS MARRV SIEOM NYYEH LPFUD MPROI RSNEI TYUAI IILLO OLRNG

CEIGF DWCNA MTRFB IIATE AAIET PLINR PCFHL COTYH EATEO NACY.

Another way to handle this same message is with a square block 10 × 10. In that event, since we have only 99 letters, we must add another letter at the end of our message to make 100. This null, or non-significant letter is simply added to fill up the square and it must be added before enciphering the plain text. This is a common rule that is rather rigidly adhered to. Let us use the same methods of inscription and transcription as in the first two examples and the letter X as a null:

E V E R Y O F F I C
E R S H O U L D B E
F A M I L I A R W I
T H T H E P R I N C
I P L E S O F S I G
N A L C O M M U N I
C A T I O N A N D E
L E M E N T A R Y M
I L I T A R Y C R Y
P T O G R A P H Y X

The cryptogram produced is:

EEFTI NCLIP VRAHP AAELT ESMTL LTMIO RHIHE CIETG YOLES OONAR

OUIPO MNTRA FLARF MAAYP FDRIS UNRCH IBWNI NDYRY CEICG IEMYX.

Another example will be shown, taking as the text the Third Amendment to the Constitution of the United States, (Article III. of the so-called Bill of Rights.)

"NO SOLDIER SHALL IN TIME OF PEACE BE QUARTERED IN ANY HOUSE WITHOUT THE CONSENT OF THE OWNER NOR IN TIME OF WAR BUT IN A MANNER TO BE PRESCRIBED BY LAW."

120 letters are in the text and the factors of 120 are 2, 2, 2, 3, and 5. These factors yield the following possible dimensions for the rectangle:

2 × 60, 60 × 2, 3 × 40, 40 × 3, 4 × 30, 30 × 4, 5 × 24, 24 × 5,

6 × 20, 20 × 6, 8 × 15, 15 × 8, 10 × 12, and 12 × 10.

Following the rule to use the most compact figures, we will employ 10 × 12 and 12 × 10, with the following inscription and transcription routes:

a. Inscription - First case under (c), Alternate Horizontal.

b. Transcription - First case under (e), Simple Diagonal.

RECTANGLE 10 × 12

```
N O S O L D I E R S
E M I T N I L L A H
O F P E A C E B E Q
N I D E R E T R A U
A N Y H O U S E W I
O C E H T T U O H T
N S E N T O F T H E
N I R O N R E N W O
T I M E O F W A R B
E N N A M A N I T U
R T O B E P R E S C
W A L Y B D E B I R
```

RECTANGLE 12 × 10

```
N O S O L D I E R S H A
E P F O E M I T N I L L
A C E B E Q U A R T E R
E S U O H Y N A N I D E
W I T H O U T T H E C O
W O E H T F O T N E S N
N E R N O R I N T I M E
M A N I T U B R A W F O
A N N E R T O B E P R E
W A L Y B D E B I R C S
```

The cipher text for the rectangle 10 × 12 will read:

NOESM OOIFN LTPIA DNEDN OIIAE YCNEL CRHES NRLEE OHEIT SABTU

TNRIE HERST TOMNR QAEUO NENTW UWOFR OAOAI HTEFM BLTHM WAEYE

WANPB ORIRD BTEEU SBCIR.

The cryptogram for the other rectangle, 12 × 10, is:

NOESP AOFCE LOESW DEBUI WIMEO TONEI QHHEE MRTUY OHRAA SNANU

TNNNW HIRAT FOINA ALTNT ORTEL LEIHT IURYR DENNB TBECE TRODO

SIABE NMWEB EFPIO RRECS.

To reconvert into plain text the cryptograms enciphered by means of geometric route transposition, the steps employed in cryptographing the original message are simply reversed. The number of letters in the cipher text determines the size of the rectangle previously agreed upon by the correspondents.

The cryptogram at the bottom of page 39 will be the first example.

EEFTI NCLIP VRAHP AAELT ESMTL LTMIO RHIHE CIETG YOLES OONAR

OUIPO MNTRA FLARF MAAYP FDRIS UNRCH IBWNI NDYRY CEICG IEMYX.

There are 100 letters and the addressee knows that a square 10 × 10 has been used.

First Step - Inscription	Second Step - Transcription
Case 1, Simple Vertical	Case 1, Simple Horizontal
E V E R Y O F F I C	EVERY OFFICER SHOULD BE FAMILIAR
E R S H O U L D B E	
F A M I L I A R W I	WITH THE PRINCIPLES OF SIGNAL
T H T H E P R I N C	
I P L E S O F S I G	COMMUNICATION AND ELEMENTARY MILITARY
N A L C O M M U N I	
C A T I O N A N D E	CRYPTOGRAPHY X..
L E M E N T A R Y M	
I L I T A R Y C R Y	
P T O G R A P H Y X	

The final letter, X, is the null inserted to complete the rectangle, since there are only 99 letters in the plain text. Note particularly that when the plain text was enciphered the inscription was Case 1, Simple Horizontal and the transcription Case 1, Simple Vertical, while in deciphering the procedure was completely reversed and we inscribed with Simple Vertical and transcribed according to Simple Horizontal. In this particular case, and occasionally in some others, it is possible to solve the cipher by using the same routes for deciphering as for enciphering. This is due to the routes chosen.

Example 2.

NOESP AOFCE LOESW DEBUI WIMEO TOMEI QHHFE MRTUY OHRAA SNANU

TNNNW HIRAT FOINA ALTNT ORTEL LEIHT IURYR DENNB TBECE TRODO

SIABE NMWEB EFPIO RRECS.

First Step - Inscription	Second Step - Transcription
Case 1, Simple Diagonal	Case 1, Alternate Horizontal
N O S O L D I E R S H A	NO SOLDIER SHALL IN TIME
E P F O E M I T N I L L	
A C E B E Q U A R T E R	OF PEACE BE QUARTERED IN
E S U O H Y N A N I D E	
W I T H O U T T H E C O	ANY HOUSE WITHOUT THE CONSENT
W O E H T F O T N E S N	
N E R N O R I N T I M E	OF THE OWNER NOR IN TIME OF
M A N I T U B R A W F O	
A N N E R T O B E P R E	WAR BUT IN A MANNER TO BE PRESCRIBED
W A L Y B D E B I R C S	
	BY LAW.

Decrypting Geometric Route Transpositions

After deciding that the cipher used is transposition and that it is of the geometric route type, the letters of the cryptogram are counted to find a clue to the probable size of the rectangle employed. As soon as the dimensions of the rectangle are determined the solution becomes a matter of trial and elimination. The letters of the cryptogram are inscribed in the rectangle in the normal way, from left to right in consecutive rows, and the horizontal lines are read forward, then back, and alternately forward and back in an endeavor to discover the plain text. Readings are next made vertically up and down the columns, followed by a like inspection of the diagonals, and finally the spiral routes are tried, both clockwise and counter-clockwise. If these efforts are unsuccessful other routes of inscription are experimented with. Constant practice in reading backward along the rows and up and down the columns and diagonals will be very helpful. By this method groups such as the following will soon be easily recognized:

LLIW for WILL, EHT for THE, DNA for AND:

```
E          E              R       F
R  and     R   for ARE:   O  and   O   for FOR.
A          A              F       R
```

Example of Decryptment

CRYPTOGRAM:

ENTEA VMLIH YINIT IOGLO GNNTL FRTFR OFBIO TAEWI RVCOL NPAVI

EOPIS TORED RUOOI RRDMG ANUNT YTWIE PTRHI BEOLA PEFAO ADOET

ARISN TMDSW RARRD OEEAE TSEED ADNTT ASSIT ABCIN LCAHG YKTEN

STDAA PTOED.

No collateral information is available. The cipher must be tackled "cold." The cryptogram contains 160 letters. Vowels and consonants are counted.

	CRYPTOGRAM		NORMAL
	Number	Per Cent	Percentage
A E I O U	62	38.75	40
L N R S T	54	33.75	30
J K Q X Z	1	.63	2
B C D F G H M P V W Y	43	26.87	28

A quick check of the individual letters shows that their recurrences are close to the normal. With this fact, and the total percentages for the four groups of vowels and consonants each well within the variation range of 5 per cent, we are justified in deciding that a transposition system has been employed.

Assuming that a geometric route cipher is involved, we check the factors of 160, the number of letters in the cryptogram. They are 2, 2, 2, 2, 2, and 5, providing a choice of the following rectangles:

2 × 80, 80 × 2, 4 × 40, 40 × 4, 5 × 32, 32 × 5, 8 × 20

20 × 8, 10 × 16, and 16 × 10.

The two most compact blocks are 10 × 16 and 16 × 10. Both horizontal and vertical inscriptions are tried and the number of vowels counted in each row and column.

Rectangle 10 × 16

Horizontal

	1	2	3	4	5	6	7	8	9	10	Vowels
1	E	N	T	E	A	V	M	L	I	H	4
2	Y	I	N	I	T	I	O	G	L	O	5
3	G	N	N	T	L	F	R	T	F	R	-
4	O	F	B	I	O	T	A	E	W	I	6
5	R	V	C	O	L	N	P	A	V	I	3
6	E	O	P	I	S	T	O	R	E	D	5
7	R	U	O	O	I	R	R	D	M	C	4
8	A	N	U	N	T	Y	T	W	I	E	4
9	P	T	R	H	I	B	E	O	L	A	4
10	P	E	F	A	O	A	D	O	E	T	6
11	A	R	I	S	N	T	M	D	S	W	2
12	R	A	R	R	D	O	E	E	A	E	6
13	T	S	E	E	D	A	D	N	T	T	3
14	A	S	S	I	T	A	B	C	I	N	4
15	L	C	A	H	G	Y	K	T	E	N	2
16	S	T	D	A	A	P	T	O	E	D	4
Vowels	6	5	5	1Ø	6	5	5	6	8	6	62

(a)

Vertical

	1	2	3	4	5	6	7	8	9	10	Vowels
1	E	O	B	V	I	P	D	R	T	G	3
2	N	G	I	I	R	T	O	R	T	Y	3
3	T	L	O	E	R	R	E	D	A	K	4
4	E	O	T	O	D	H	T	O	S	T	4
5	A	G	A	P	M	I	A	E	S	E	6
6	V	N	E	I	G	B	R	E	I	N	4
7	M	N	W	S	A	E	I	A	T	S	4
8	L	T	I	T	N	O	S	E	A	T	4
9	I	L	R	O	U	L	N	T	B	D	3
10	H	F	V	R	N	A	T	S	C	A	2
11	Y	R	C	E	T	P	M	E	I	A	4
12	I	T	O	D	Y	E	D	E	N	P	4
13	N	F	L	R	T	F	S	D	L	T	-
14	I	R	N	U	W	A	W	A	C	O	5
15	T	O	P	O	I	O	R	D	A	E	6
16	I	F	A	O	E	A	A	N	H	D	6
Vowels	7	3	7	9	5	8	5	8	5	5	62

(b)

FIGURE 8

The normal number of vowels A E I O and U should total 4 for each row and 6.4 for each column, while all should add up to 64.

Although horizontal reading from left to right shows good digraphs and two excellent trigraphs, TIO in the second row and RED at the end of row 6, this

evidence is insufficient. Reading down the columns of (a) in Figure 8 yields interesting results. Every one of the 10 columns discloses portions of plain text too numerous to be merely a matter of co-incidence.

	1	2	3	4	5	6	7	8	9	10
1										
2		I		I						
3		N		T		F			F	
4		F		I						
5				O		N	P	A	V	I
6		O		I	S	T	O	R	E	D
7		U	O	O	I	R	R	D	M	G
8	A	N	U	N	T	Y	T	W	I	E
9	P	T	R	H	I	B	E	O	L	A
10	P	E	F	A	O	A	D	O	E	T
11	A	R	I	S	N	T		D	S	W
12	R		R	R				E		E
13				E						
14										

Column 1 shows APPAR, which may be a part of APPARENT or APPARENTLY. In Column 2 are the INF of INFANTRY and -OUNTER from COUNTER- or ENCOUNTER. OUR and the FIR- of FIRST are in Column 3, while in Column 4 we see ITIO and -ION HAS RE. Column 5 yields SITION, and Column 6 discloses F-NTRY BAT-, the latter suggesting INFANTRY BATTALION. In Column 7 we have -PORTED, which may be a portion of REPORTED or SUPPORTED. Column 8 gives us ARD WOOD E- and Column 9 - F-VE MILES. Lastly, Column 10 shows -IDGE AT WE-. We seem to be on the right track and the message is obviously military in character. Let us try out the block 16 columns wide and 10 rows deep with both Horizontal and Vertical inscriptions and see if we get further clues.

Horizontal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Vowels
1	E	N	T	E	A	V	M	L	I	H	Y	I	N	I	T	I	7
2	O	G	L	O	G	N	N	T	R	F	R	T	F	R	O	F	3
3	B	I	O	T	A	E	W	I	R	V	C	O	L	N	P	A	7
4	V	I	E	O	P	I•	S	T	O	R	E	D	R	U	O	O	9
5	I	R	R	D	M	G	A	N	U	N	T	Y	T	W	I	E	5
6	P	T	R	H	I	B	E	O	L	A	P	E	F	A	O	A	8
7	D	O	E	T	A	R	I	S	N	T	M	D	S	W	R	A	5
8	R	R	D	O	E	E	A	E	T	S	E	E	D	A	D	N	8
9	R	R	D	O	E	E	A	E	T	S	E	E	D	A	D	N	5
10	G	Y	K	T	E	N	S	T	D	A	A	P	T	O	E	D	5
Vowels	3	3	4	4	6	4	4	4	3	2	4	4	-	5	6	6	62

Average number of vowels in each row should be 6.4 and in each column 4, with a total of 64. A survey of this inscription reveals that the possibilities are not nearly so good as the previous test. Now for the Vertical case.

VERTICAL

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Vowels
1	E	Y	G	O	R	E	R	A	P	P	A	R	T	A	L	S	6
2	N	I	N	F	V	O	U	N	T	E	R	A	S	S	C	T	5
3	T	N	N	B	C	P	O	U	R	F	I	R	E	S	A	D	5
4	E	I	T	I	O	I	O	N	H	A	S	R	E	I	H	A	10
5	A	T	L	O	L	S	I	T	I	O	N	D	D	T	G	A	6
6	V	I	F	T	N	T	R	Y	B	A	T	O	A	A	Y	P	5
7	M	O	R	A	P	O	R	T	E	D	M	E	D	B	K	T	5
8	L	G	T	E	A	R	D	W	O	O	D	E	N	C	T	O	6
9	I	L	F	W	V	E	M	I	L	E	S	A	T	I	E	E	8
10	H	O	R	I	I	D	G	E	A	T	W	E	T	N	N	D	6
Vowels	4	5	-	6	3	5	4	4	4	6	2	6	3	4	2	4	62

FIGURE 9

Examination reveals that diagonals and horizontals yield not only the words and portions of words shown in the vertical columns of Figure 8, (a), but additional letters that give better sense to the extracts we have found. The diagonals incline us to believe that a trial of this rectangle with an inscription according to one of the routes under (e) on Page 37 would be worth while. First we will attempt it by the first case, Simple Diagonal, beginning at the upper left-hand corner.

E N E M Y I N F A N T R Y B A T
T A L I O N R E P O R T E D M O
V I N G T O W A R D W O O D E D
H I L L F I V E M I L E S E A S
T O F B R I D G E A T W A D I N
G R I V E R A P P A R E N T L Y
T O C O U N T E R A T T A C K S
T O P O U R F I R S T B A T T A
L I O N H A S R E A C H E D P O
S I T I O N D E S I G N A T E D

The plain text message is revealed as each row is read in the usual way, from left to right.

The complete plain text is:

ENEMY INFANTRY BATTALION REPORTED MOVING TOWARD WOODED HILL
FIVE MILES EAST OF WADING RIVER APPARENTLY TO COUNTER-ATTACK
STOP OUR FIRST BATTALION HAS REACHED POSITION DESIGNATED.

Reversed, or "mirror writing," is a special case of route transposition, although it comes under the general classification of case (a), 2, Simple Horizontal, as shown on page 37.

The message, "COME UP AND SEE ME SOME TIME," may be written in several different ways.

a. Completely reversed, normal word-lengths retained:

EMIT EMOS EM EES DNA PU EMOC.

b. Completely reversed and divided into 5-letter groups:

EMITE MOSEM EESDN APUEM OC.

c. Completely reversed, nulls added and divided into 5-letter groups:

EMITE MOSEM EESDN APUEM OCNEL.

d. Each word reversed, nulls added, and divided into 5-letter groups:

EMOCU PUERB DNAYX EESEN EMZOF EMOSJ EMITH

There are other variations of this method. The underlined letters are nulls. The well-known "Rail-Fence" cipher belongs to the route transposition class in case 1 of (b), Simple Vertical, as to inscription, and classed as case 1 of (a), Simple Horizontal, for transcription. With the message, COME UP AND SEE ME SOME TIME, the inscription is:

C M U A D E M S M T M
O E P N S E E O E I E

and the cipher text, as prepared for telegraphic transmission:

CMUAD EMSMT MOEPN SEEOE IE.

It is usually written with the two rows staggered, thus:

C M U A D E M S M T M
O E P N S E E O E I E .

If diagonal lines are drawn between the letters the appearance of these is similar to the old rail fences often seen on farms.

Chapter VIII

COLUMNAR TRANSPOSITION CIPHERS, WITH COMPLETELY FILLED FIGURES

In addition to the route transposition ciphers described in Chapter VII, another method frequently employed is columnar transposition. A geometric figure is commonly used and the most popular design is the rectangle. Columnar transpositions are divided into two types: one in which all of the small squares or cells of the block are filled in, even though nulls must be added to do this; the second in which some of the cells are left blank. In the latter type some of the columns will have fewer rows than the others. The first is called a completely filled design and the second an incomplete one. Although other geometric patterns may be used, this chapter will deal only with rectangles.

The letters of the plain text are inscribed in the figure in the usual way from left to right in successive rows. Transcription is by columns to form the cryptogram. The order in which the columns are taken is usually determined by a derived numerical key. In the example to follow the key-word is CUMBERLAND and the derived numerical key is 3-10-7-2-5-9-6-1-8-4. The plain text is: "C.O. FIRST INF. ENCOUNTERED BLUE INFANTRY ESTIMATED TWO COMPANIES SOUTH OF TOONERVILLE NEAR HALLS FARM ENEMY RETREATING WESTWARD OUR LOSSES SLIGHT AM PURSUING. FORD, MAJ." It consists of 140 letters and the numerical key indicates 10 columns, so there will be 14 rows. The numerical key is written at the top.

	3	1Ø	7	2	5	9	6	1	8	4
1	C	O	F	I	R	S	T	I	N	F
2	E	N	C	O	U	N	T	E	R	E
3	D	B	L	U	E	I	N	F	A	N
4	T	R	Y	E	S	T	I	M	A	T
5	E	D	T	W	O	C	O	M	P	A
6	N	I	E	S	S	O	U	T	H	O
7	F	T	O	O	N	E	R	V	I	L
8	L	E	N	E	A	R	H	A	L	L
9	S	F	A	R	M	E	N	E	M	Y
1Ø	R	E	T	R	E	A	T	I	N	G
11	W	E	S	T	W	A	R	D	O	U
12	R	L	O	S	S	E	S	S	L	I
13	G	H	T	A	M	P	U	R	S	U
14	I	N	G	F	O	R	D	M	A	J

(FIGURE 10)

Beginning with column 1, we transcribe the first 14 letters:

I E F M M T V A E I D S R M

and follow this with the letters from column 2:

I O U E W S O E R R T S A F .

Then we continue with the letters in columns 3, 4, and 5, and so on until the cryptogram is complete:

IEFMM TVAEI DSRMI OUEWS OERRT SAFCE DTENF LSRWR GIFEN TAOLL

YGUIU JRUES OSNAM EWSMO TTNIO URHNT RSUDF CLYTE ONATS OTGNR

AAPHI LMNOL SASNI TCOER EAAEP RONBR DITEF EELHN.

To decryptograph the above message, the addressee consults his key-word and obtains the derived numerical key, which indicates 10 columns. Then he counts the letters in the cryptogram (140) and divides this by 10 (the number of columns). This gives him 14, which is the number of rows, and his rectangle is, therefore, 10 columns wide by 14 rows deep. He proceeds to fill in the letters of the secret text by writing down the 14 letters from column 1, placing them in the column having this number at the top. Then he writes the next 14 letters in the column numbered 2, continuing until he has filled the rectangle, see Figure 11, below.

3	10	7	2	5	9	6	1	8	4
							I		
							E		
							F		
							M		
							M		
							T		
							V		
							A		
							E		
							I		
							D		
							S		
							R		
							M		

(a)

3	10	7	2	5	9	6	1	8	4
			I				I		
			O				E		
			U				F		
			E				M		
			W				M		
			S				T		
			O				V		
			E				A		
			R				E		
			R				I		
			T				D		
			S				S		
			A				R		
			F				M		

(b)

FIGURE 11

The plain text is read row by row in the usual way from top to bottom.

The route of inscription is not limited to the Simple Horizontal as shown in the example, but may be any one of the others given in Chapter VII.

In addition to the numerical key method of transcription, the order in which the columns are taken may be in accordance with special agreements between the correspondents. One of these variations is the route cipher with words, used during the War between the States by the Federal Army, a specimen of which is given below:

Message: "WHAT ARE THE ENEMY DOING I SHALL MOVE ALL MY ARTILLERY TONIGHT AND SHALL ATTACK AT DAYLIGHT SEND ME ONE THOUSAND CAVALRY FOUR THOUSAND INFANTRY AND ALL THE PROVISIONS YOU CAN SPARE."

A rectangle 6 × 6 is used. The route for inscription is: Up column 1, Down column 3, Up column 2, Down column 5, Up column 6, and Down Column 4.

	1	2	3	4	5	6
1	I	SEND	SHALL	CAN	ME	YOU
2	DOING	DAYLIGHT	MOVE	SPARE	ONE	PROVISIONS
3	ENEMY	AT	ALL	JONES	THOUSAND	THE
4	THE	ATTACK	MY	IS	CAVALRY	ALL
5	ARE	SHALL	ARTILLERY	VERY	FOUR	AND
6	WHAT	AND	TONIGHT	SICK	THOUSAND	INFANTRY

Since the plain text contains only 32 words, while the square has 36 divisions, the four words, JONES IS VERY SICK, are added as nulls to fill up column 4. Transcription is by rows from top to bottom, starting with Row 1, and the cipher text reads:

" I SEND SHALL CAN ME YOU DOING DAYLIGHT MOVE SPARE ONE PROVISIONS ENEMY AT ALL JONES THOUSAND THE THE ATTACK MY IS CAVALRY ALL ARE SHALL ARTILLERY VERY FOUR AND WHAT AND TONIGHT SICK THOUSAND INFANTRY."

The underlined words are the nulls.

Decrypting the Cipher

To solve by cryptanalysis a message enciphered by columnar transposition employing a completely filled rectangle, it is only necessary to determine the number of columns and rows (the dimensions) of the figure used, inscribe the cryptogram vertically in the columns from left to right, and then anagram the columns. It is wise to make a carbon copy of the inscription, using quadrille ruled paper for both original and carbon copy.

Taking the cryptogram on Page 48,

IEFMM TVAEI DSRMI OUEWS OERRT SAFCE DTENF LSRWR GIFEN TAOLL

YGUIU JRUES OSNAM EWSMO TINIO URHNT RSUDF CLYTE ONATS OTGNR

AAPHI LMNOL SASNI TCOER EAAEP RONBR DITEF EELHN,

there are 140 letters. The factors are 2, 2, 5, and 7, yielding two fairly compact rectangles, 10 × 14 and 14 × 10. The probability that a ten-letter key-word and derived numerical key has been used is better than the case for a 14-letter key-word, 10-letter words being more numerous and more easily memorized. The inscription will be

	1	2	3	4	5	6	7	8	9	1Ø
1	I	I	C	F	R	T	F	N	S	O
2	E	O	E	E	U	T	C	R	N	N
3	F	U	D	N	E	N	L	A	I	B
4	M	E	T	T	S	I	Y	A	T	R
5	M	W	E	A	O	O	T	P	C	D
6	T	S	N	O	S	U	E	H	O	I
7	V	O	F	L	N	R	O	I	E	T
8	A	E	L	L	A	H	N	L	R	E
9	E	R	S	Y	M	N	A	M	E	F
1Ø	I	R	R	G	E	T	T	N	A	E
11	D	T	W	U	W	R	S	O	A	E
12	S	S	R	I	S	S	O	L	E	L
13	R	A	G	U	M	U	T	S	P	H
14	M	F	I	J	O	D	G	A	R	N

FIGURE 12

The collateral information is scant. It is known that the message was exchanged by two units of enemy troops consisting of not more than a regiment of infantry.

Examination of the rows for likely occurrences of the frequent reversed di-

graphs ER and RE and the frequent trigraph ERE provides one logical method of attack, and since the message may refer to our own forces, we will also look for the word ENEMY, or rather - the letters forming this word. The 2nd, 8th, 9th, and 10th rows are the only ones containing an R and at least two E's, but in the ninth row we find all the letters of the word ENEMY, which are in columns 1, 2, 4, 5, 8, and 9. Now we will take our carbon copy of the rectangle and cut the columns apart so that we will have ten strips of vertical letters, one for each column, and numbered at the top. We can now anagram the columns. In the second row R is in column 8, with the three E's in columns 1, 3, and 4. We will place columns 1, 3, and 4 alongside of column 8 in turn to form the digraph ER in the second row and then glance down the columns for other likely digraphs, as in (a) of Figure 13, below.

	1 8	3 8	4 8
1	I N	C N	F N
2	E R	E R	E R
3	F A	D A	N A
4	M A	T A	T A
5	M P	E P	A P
6	T H	N H	O H
7	V I	F I	L I
8	A L	L L	L L
9	E M	S M	Y M
10	I N	R N	G N
11	D O	W O	U O
12	S L	R L	I L
13	R S	G S	U S
14	M A	I A	J A

(a)

	1 8 3	1 8 4
1	I N C	I N F
2	E R E	E R E
3	F A D	F A N
4	M A T	M A T
5	M P E	M P A
6	T H N	T H O
7	V I F	V I L
8	A L L	A L L
9	E M S	E M Y
10	I N R	I N G
11	D O W	D O U
12	S L R	S L I
13	R S G	R S U
14	M A I	M A J

(b)

	9 6 1 8 4
1	S T I N F
2	N T E R E
3	I N F A N
4	T I M A T
5	C O M P A
6	O U T H O
7	E R V I L
8	R H A L L
9	E N E M Y
10	A T I N G
11	A R D O U
12	E S S L I
13	P U R S U
14	R D M A J

(c)

FIGURE 13

The first pair, columns 1 and 8, gives us better digraphs, especially since the EM of ENEMY appears in row 9. Also, in row 9 of the combination of the 4th and 8th columns there is YM, which is MY reversed. A survey of these two pairs of columns leads us to select one and eight as best and to place column 4 to the right of column 8, trying also column 3 in this position as in (b), above. The trigraphs in columns 1, 8, and 4 are much better than those in the set-up 1, 8, and 3. Note the MAJ in the 14th row. This may be the abbreviation for or part of the word

Major. Referring again to the 9th row, in (b) we now have the EMY of ENEMY, so we look for EN in that row, finding the E in column 9 and the N in column 6. The 6th column is now placed to the left of column 1 and the 9th on the left of 6. We have at this moment five of the columns placed in the order 9-6-1-8-4, as shown in (c) of Figure 13. Note the word ENEMY in row 9. This is correct and to confirm our suppositions thus far, it is only necessary to observe in the 3rd row INFAN of INFANTRY, in the 5th COMPA of COMPANY, in the 13th PURSU of PURSUE, PURSUIT or PURSUING. Likewise, in Row 2 is found NTERE, which may be part of ENCOUNTERED. By the trial and error method we now endeavor to fit the other five columns to the five we have already aligned. The completed rectangle is:

	3	1Ø	7	2	5	9	6	1	8	4
1	C	O	F	I	R	S	T	I	N	F
2	E	N	C	O	U	N	T	E	R	E
3	D	B	L	U	E	I	N	F	A	N
4	T	R	Y	E	S	T	I	M	A	T
5	E	D	T	W	O	C	O	M	P	A
6	N	I	E	S	S	O	U	T	H	O
7	F	T	O	O	N	E	R	V	I	L
8	L	E	N	E	A	R	H	A	L	L
9	S	F	A	R	M	E	N	E	M	Y
1Ø	R	E	T	R	E	A	T	I	N	G
11	W	E	S	T	W	A	R	D	O	U
12	R	L	O	S	S	E	S	S	L	I
13	G	H	T	A	M	P	U	R	S	U
14	I	N	G	F	O	R	D	M	A	J

FIGURE 14

The numerical key is 3-1Ø-7-2-5-9-6-1-8-4. It is not necessary for us to attempt to guess the literal key. It may be a 1Ø-letter word or a phrase or even two shorter words. For our needs the numerical key is enough. In this problem both frequential analysis of digraphs and trigraphs and the "probable word" method of finding the word ENEMY were used together.

Chapter IX

COLUMNAR TRANSPOSITION CIPHERS, WITH INCOMPLETELY FILLED FIGURES

Instead of using a completely filled design, incomplete figures are often used. A derived numerical key usually determines the number of columns and the number of letters in the plain text divided by the length of the key fixes the number of rows. In this type of transposition some of the columns being shorter will contain fewer letters that the other columns.

Taking the key-word BUCKINGHAM, which yields the numerical key 2-1Ø-3-7-6-9-4-5-1-8, we will proceed to encipher the following message:

"A WELL REGULATED MILITIA BEING NECESSARY TO THE SECURITY OF A FREE STATE THE RIGHT OF THE PEOPLE TO KEEP AND BEAR ARMS SHALL NOT BE INFRINGED."

	2	1Ø	3	7	6	9	4	5	1	8
1	A	W	E	L	L	R	E	G	U	L
2	A	T	E	D	M	I	L	I	T	I
3	A	B	E	I	N	G	N	E	C	E
4	S	S	A	R	Y	T	O	T	H	E
5	S	E	C	U	R	I	T	Y	O	F
6	A	F	R	E	E	S	T	A	T	E
7	T	H	E	R	I	G	H	T	O	F
8	T	H	E	P	E	O	P	L	E	T
9	O	K	E	E	P	A	N	D	B	E
1Ø	A	R	A	R	M	S	S	H	A	L
11	L	N	O	T	B	E	I	N	F	R
12	I	N	G	E	D					

FIGURE 15

Columns 2, 1Ø, 3, 7, and 6 each contain 12 letters, while columns 9, 4, 5, 1, and 8 have only 11 letters each. Row 12 has only 5 letters, but all the others have 1Ø each.

In transcribing the cryptogram for transmission we follow the procedure used for completely filled rectangles (as described in Chapter VIII.), beginning with the letters in the column headed 1, followed by those in the columns numbered 2, 3, 4, in regular order, continuing until the entire communication is written out.

UTCHO TOEBA FAAAS SATTO ALIEE EACRE EEAOG ELNOT THPNS IGIET

YATLD HNLMN YREIE PMBDL DIRUE RPERT ELIEE FEETE LRRIG TISGO

ASETD BSEFH HKRNN.

Steps in the deciphering of such cryptograms are similar to those for decrypto-graphing ciphers of the completely filled block type. The addressee takes the key-word agreed upon with the sender, derives the numerical key and arranges his col-umns accordingly. Then he writes the letters of the cryptogram in the columns in order according to the numbers at the tops of the columns, completing column 1, then column 2, and so on, until the whole of the cipher text has been filled in. Entries in the first 3 columns are:

Key-word:	B	U	C	K	I	N	G	H	A	M
Numerical Key:	2	1Ø	3	7	6	9	4	5	1	8
1	A		E						U	
2	A		E						T	
3	A		E						C	
4	S		A						H	
5	S		C						O	
6	A		R						T	
7	T		E						O	
8	T		E						E	
9	O		E						B	
1Ø	A		A						A	
11	L		O						F	
12	I		G							

FIGURE 16

When all columns are filled in the message is read by horizontal rows.

Cryptanalysis of Incomplete Rectangles

The following cryptogram has been received for solution:

OATRF SOISW PDEEH IPWAR ENEOP ITTEL URDES AITCE EFOEC BAFMN

TETSM ETATF DMRNI HO.

No collateral information is available, but the message is suspected of containing subject matter concerning the freedom of the press. The cryptogram contains 67 letters and is believed to be enciphered with an incomplete rectangle. Blocks 7 × 1Ø and 1Ø by 7 are the nearest in size to take 67 letters. We will try the 7 × 1Ø rectangle first.

Inscription

	1	2	3	4	5	6	7
1	O	P	E	U	E	N	T
2	A	D	N	R	F	T	F
3	T	E	E	D	O	E	D
4	R	E	O	E	E	T	M
5	F	H	P	S	C	S	R
6	S	I	I	A	B	M	N
7	O	P	T	I	A	E	I
8	L	W	T	T	F	T	H
9	S	A	E	C	M	A	O
1Ø	W	R	L	E			

FIGURE 17

We know that the first four columns have 1Ø letters each and the last three have only 9 each. The cryptogram was set up on quadrille-ruled paper as in the figure above and a carbon copy made. Now we take the carbon copy and cut the figure into vertical strips so that each strip contains one column with its number at the top, so that the strips can be re-arranged in sequence and slid against each other.

Let us look at the various rows and see if any of them contain the letters of the word PRESS. Row 5 contains four of the letters, P in column 3, R in column 7, and the two S's in columns 4 and 6. This row also contains the letter H. We will place column 3 to the left of column 7. This anagramming should give an answer to the question of whether the remaining digraphs that will be in these two columns are probable and what might reasonably be expected to appear in so short a message.

Since H of THE in the phrase FREEDOM OF THE PRESS has only one letter, E, between it and the P of PRESS, and H in this row is in column 2, we will place column 2 to the left of 3 but allow one space between them to insert the column that may contain the E to follow the H. Columns 4 and 6 will be lined up to the right of 7 and our partial block will look like Figure 18, below.

	2	3	7	4	6
1	P	E	T	U	N
2	D	N	F	R	T
3	E	E	D	D	E
4	E	O	M	E	T
5	H	P	R	S	S
6	I	I	N	A	M
7	P	T	I	I	E
8	W	T	H	T	T
9	A	E	O	C	A
1Ø	R	L		E	

FIGURE 18

	2	3	7	6
1	P	4	E	T	N
2	D	U	N	F	T
3	E	R	E	D	E
4	E	D	O	M	T
5	H	E	P	R	S
6	I	S	I	N	M
7	P	A	T	I	E
8	W	I	T	H	T
9	A	T	E	O	A
1Ø	R	C	L		
		E			

FIGURE 19

	1	2	3	7	5	6
1	P	4	E	T	A	N
2	D	U	N	F	E	T
3	E	R	E	D	F	E
4	E	D	O	M	O	T
5	H	E	P	R	E	S
6	I	S	I	N	C	M
7	P	A	T	I	B	E
8	W	I	T	H	A	T
9	A	T	E	O	F	A
1Ø	R	C	L		M	
		E				

FIGURE 20

The digraphs in columns 3,7, are all good, but the combinations in 4,6, are poor.

In row 4 we note in columns 3,7, the digraph OM, probably part of FREEDOM. Now let us look closely at column 4. We need a D with an E below it so that the D will fit alongside of OM in 3,7, and that the E beneath the D will fit in between the H and P of row 5. Do we find a possibility in column 4? Yes, so let us place column 4 between columns 2 and 3, but drop it down one row below, since the D and E of column 4 are in rows 3 and 4, while we need them in rows 4 and 5. Our block now appears as in Figure 19 above. Thus far the combinations of letters are all good. Columns 1 and 5 have not been used, although they were naturally checked when we fitted the D to the OM. Column 5 contains an E in row 4 that will fit for the E after the PR of the suspected word PRESS. Column 1 does not contain an E, but has an S in the 6th row that might do for the first S of PRESS. We now put column 5 to the right of 7, column 1 to its right and to the left of 6. We must drop column 5 one row as we did with column 4 and raise column 1 one row. The block now appears as in Figure 20 and the 2nd to 9th rows, inclusive, make good sense. The order of the columns is 2-4-3-7-5-1-6. Three letters in row 1Ø of

Figure 2Ø, C, L, and M, seem out of place, but the R is probably in its correct sequence to follow the WA at the end of row 9, to form the word WAR. The E below the C in column 4 is also out of place.

What shall we do with the four letters C, E, L, and M? If we bring up the C and E from the base of column 4 and place them above the other letters in column 5, put the L from the bottom of column 3 at the right of the first P in column 4, and the M from the base of column 5 above the N, we will have a block like that in Figure 21, and the entire plain text can be read. Now let us re-write the entire cryptogram in a 7 × 10 block, but making the order of the columns 5-1-6-2-4-3-7. Why do we change the sequence of the columns? Because we have a clear indication that this is the proper order. The first seven letters of the clear text, COMPLET, as in Figure 21, appear in the columns in the sequence 5-1-6-2-4-3-7. Figure 22 shows the correct arrangement by which the message was enciphered.

2	4	3	7	5	1	6
				C	O	M
P	L	E	T	E	A	N
D	U	N	F	E	T	T
E	R	E	D	F	R	E
E	D	O	M	O	F	T
H	E	P	R	E	S	S
I	S	I	N	C	O	M
P	A	T	I	B	L	E
W	I	T	H	A	S	T
A	T	E	O	F	W	A
R	C	L		M		
	E					

FIGURE 21

5	1	6	2	4	3	7
1	C	O	M	P	L	E	T
2	E	A	N	D	U	N	F
3	E	T	T	E	R	E	D
4	F	R	E	E	D	O	M
5	O	F	T	H	E	P	R
6	E	S	S	I	S	I	N
7	C	O	M	P	A	T	I
8	B	L	E	W	I	T	H
9	A	S	T	A	T	E	O
1Ø	F	W	A	R			

FIGURE 22

The plain text - "COMPLETE AND UNFETTERED FREEDOM OF THE PRESS IS INCOMPATIBLE WITH A STATE OF WAR," is a quotation from Bronsart von Schellendorf, who was a writer on military affairs.

Chapter X

SIMPLE SUBSTITUTION, MONOALPHABETIC SYSTEMS

The simplest form of substitution is one in which a single cipher alphabet is used throughout an entire message or series of messages. This is known as monoalphabetic substitution. With a single cipher alphabet each letter is in the plain text is always represented by the same character in the cipher text; for example, if "E" in the clear is replaced in the cryptogram by "K" the first time the letter "E" appears in the plain text, then every time thereafter that "E" occurs in the clear it is replaced by "K" in the cipher text. The equation showing this is - E_p equals K_c ; or K_c equals E_p .

Augustus Caesar simply replaced each letter in the clear by the letter that followed it in the normal alphabet for the Latin Language. Thus, the message, REPORT AT ONCE TO YOUR COMMANDING OFFICER, would be enciphered:

Plain - REPORT AT ONCE TO YOUR COMMANDING OFFICER } According to our own Alphabet in English.

Cipher - SFQPSU BU PODF UP ZPVS DPNNBOEJOH PGGJDFS

Note that the grouping is in normal word lengths.

Let us encipher a message, using a direct standard alphabet commencing with the letter N. The enciphering alphabet is:

Plain component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Cipher Component - N O P Q R S T U V W X Y Z A B C D E F G H I J K L M

Message: REGIMENTAL COMMAND POST WILL BE IN FARM HOUSE FOUR HUNDRED YARDS WEST OF JONESVILLE.

For encipherment and telegraphic transmission we will use 5-letter groups.

Plain - REGIM ENTAL COMMA NDPOS TWILL BEINF ARMHO USEFO URHUN

Cipher - ERTVZ RAGNY PBZZN AQCBF GJVYY ORVAS NEZUB HFRSB HEUHA

Plain - DREDY ARDSW ESTOF JONES VILLE

Cipher - QERQL NEQFJ RFGBS WBARF IVYYR

It should be carefully noted that the cipher alphabet is reciprocal. A_p equals N_c, A_c equals N_p, B_p equals O_c, and B_c equals O_p, etc. The deciphering alphabet is identical with the enciphering alphabet, thus either one may be used for both enciphering and deciphering.

Converting the cryptogram back into the clear text is simply the reverse of putting it into cipher.

Decryptment

The following cryptogram, containing 216 letters, has been received for solution. No collateral information is available, but the language suspected is English.

FDUHV KRXOG EHHAH UFLVH GWKDW RQOBR QHHGL WLRQR IDFRG HRIDS
DUWLF XODUF ODVVR URQOB RQHFL SKHUN HBLVE HLQJX VHGZL WKLQD
XQLWD WRQHW LPHZK HQDFR GHLVU HSODF HGEBD QHZFR GHRUD QHZHG
LWLRQ WKHUH SODFH GFRGH ZLOOE HGHVW URBHG EBEXU QLQJZ QOHVV
RWKHU ZLVHR UGHUH G.

Only a rapid glance is needed to spot the high frequency of the five rare consonants. They total 43, approximately 20% of the total and ten times the normal.

The next step is to recopy the message on quadrille-ruled paper having four or five squares to the inch. It is copied in the five-letter groups. Indeed, all cryptograms should be copied with the same grouping that occurs in the received document. Sometimes it may be desirable to recopy in five letter groups to facilitate analysis and the preparation of frequency charts, tables of digraphs and trigraphs, and other data.

We will now prepare a trigraphic frequency chart, since this will not only give the trigraphs, but also the single frequency of individual letters, in addition to which a list of digraphs may easily be prepared from it. The drawing up of a frequency chart of this kind will often save much time in the various steps involved in cryptanalysis and in the checking of the probable identifications of letters. The complete chart is shown on the next page.

TRIGRAPH FREQUENCY No._____

A	H H	1
B	OOHERE RRLDHE	6
C		Ø
D	FKIISOOQWQOBUO UWFSUUVXWFFQQF	14
E	GVGOGB HHBHBX	6
F	-UDLUHDDZDG DLRXOLRHRHR	11
G	OHHRHRHRHHRHHUH EWLHZHEHLFHHEH-	15
H	UEHAVQHGQKNEVQPKGUFQGQZKUFGEGBOKVGU VHAUGHGRFUBLGWZQLSGZRZGUSGZGVGVURUG	35
I	RR DD	2
J	QQ XX	2
K	VWSWZWW RDHLHHH	7
L	FGWWFBHZKQWHGWZQZ VWRFSVQWQWPVWROQV	17
M		Ø
N	U H	1
O	XQXFQSSLOQ GBDDBDDOEH	1Ø
P	L H	1
Q	RRRRRLLXRHDDRULZ OHROHJDLHDHHWLJO	16
R	KWBLQFHVUBWFFHLFUVH XQQQIGIUQQQGGUQGBWU	19
S	DLHH DKOO	4
T		Ø
U	DHDDRHVRHWXHRH HFWFRNHDHRQZGH	14
V	HLDVLXLHHVL KHVREHUWVRH	11
W	GDLULLDHLQVR KRLLKDRLLKUK	12
X	RFJDEJ OOVQUQ	6
Y		Ø
Z	GHHHHU LKFHLL	6
		216

Our chart shows marked crests and troughs and from its appearance it is evident that we have to deal with a monoalphabetic or single substitution cipher. Four letters, C, M, T, and Y, do not occur in the cryptogram. These four "blanks" may represent four of the rare consonants, J, K, Q, X, and Z. The frequency table reveals that H appears 35 times in a total of 216 letters, or 16.2%. H occurs twice as a doublet, HH, and forms digraphs with 17 different other letters. It is preceded by 15 and followed by 12 different letters. Although its percentage - 16.2 - is somewhat high for E with a normal proportion of 13%, we may assume that H represents E, because it is by far the most frequent letter in the cryptogram and combines readily with so many others. H is followed in order by R - 19 times, L - 17, Q - 16, and G - 15 times. Let us write the cryptogram again in 5-letter groups and place the letter E under H wherever the latter occurs in the message.

```
FDUHV KRXOG EHHAH UFLVH GWKDW RQOBR QHHGL WLRQR IDFRG HRIDS
  E          EE E    E                   EE             E

DUWLF XODUF ODVVR URQOB RQHFL SKHUN HBLVE HLQJX VHGZL WKLQD
                        E    E    E          E

XQLWD WRQHW LPHZK HQDFR GHLVU HSODF HGEBD QHZFR GHRUD QHZHG
      E    E    E    E    E    E    E    E    E E

LWLRQ WKHUH SODFH GFRGH ZLOOE HGHVW URBHG EBEXU QLQJX QOHVV
        E E    E    E         E E    E                   E

RWKHU ZLVHR UGHUH G.
  E    E    E E
```

We will now list the digraphs and trigraphs containing H and occurring more than once and see if they will help us in confirming the value of H as E and identifying other plain text letter equivalents.

DIGRAPHS				TRIGRAPHS			
HG -	11111	1111	- 9	RGH -	1111		- 4
GH -	11111	1	- 6	KHU -	111		- 3
HU -	11111		- 5	RQH -	111		- 3
HZ -	1111		- 4	DFH -	11		- 2
KH -	1111		- 4	FHG -	11		- 2
QH -	1111		- 4	GHR -	11		- 2
UH -	1111		- 4	HGE -	11		- 2
EH -	111		- 3	HGL -	11		- 2
HR -	111		- 3	HRU -	11		- 2
HV -	111		- 3	HUH -	11		- 2
VH -	111		- 3	LVH -	11		- 2
FH -	11		- 2	QHZ -	11		- 2
HH -	11		- 2	UHS -	11		- 2
				WKH -	11		- 2

Doubled letters occurring more than once are HH and VV, both of which appear twice. One other letter is doubled and occurs only once - OO.

The most frequent reciprocal pairs of letters in English are ER and RE, next come DE and ED, followed by ES and SE. If HG represents ER and GH represents RE, then we must check two things. First: Is the frequency of G in our cryptogram close to the normal for R ? G occurs 15 times, while R in literary text has a frequency of 12 in 200 letters and in telegraphic text one of 14. This is favorable. Second: R combines in English with all of the vowels and with many of the consonants - Do we find in the cryptogram that G combines readily with a fairly large number of the other letters? Four letters precede and seven letters follow it. Possibly G stands for R, but before assigning this value it is desirable to check a little more definitely on other possibilities. A frequent trigraph is ERE and if H_c equals E_p we should find a group H-H. Three such groups occur, HUH twice and HAH and HGH once each. What are the characteristics of U and A? Only once does A appear and it cannot equal R. The frequency of U in the message is 14. Six other letters precede and nine others follow it. The choice of an equivalent for R_p lies between U and G. What about reversed digraphs with U equal to R? HU occurs 5 times and UH 4 times. Because U combines with other letters more readily than G does, the evidence favors the supposition that U_c equals R_p. Let us tentatively assign this value:

FDUHV RE	KRXOG	EHHAH EE E	UFLVH R E	GWKDW	RQOBR	QHHGL EE	WLRQR	IDFRG	HRIDS E
DUWLF R	XODUF R	ODVVR	URQOB R	RQHFL E	SKHUN ER	HBLVE E	HLQJX E	VHGZL E	WKLQD
XQLWD	WRQHW E	LPHZK E	HQDFR E	GHLVU E R	HSODF E	HGEBD E	QHZFR E	GHRUD E R	QHZHG E E
LWLRQ	WKHUH ERE	SODFH E	GFRGH E	ZLOOE	HGHVW E E	URBHG R E	EBEXU R	QLQJX	QOHVV E
RWKHU ER	ZLVHR E	UGHUH R ERE	G.						

Now for the doubled letters, which may help in determining the cipher equivalent for S in the clear. The two most frequent doublets in English are EE and SS. In the cryptogram we have HH and VV each appearing twice. V_c has a frequency of 11, close to the normal 13. The third most frequent reciprocal pairs in the cipher are HV and VH, three times and in proper order for ES and SE in plain text. This evidence establishes the probability that V_c equals S_p and confirms further H_c as E_p.

There is another pair of reversals to consider, HG and GH. Having tentatively established ER and RE, ES and SE, another reciprocal pair that occurs very often in English is DE - ED. How will G_c equalling D_p work out? The last six letters of the cryptogram are UGHUH G and if G is the cipher equivalent of D then the plain text values of the six letters are RDERE D. Is it possible that this is part of the word ORDERED? If so, another value can be assigned, R_c equals O_p. Can this be confirmed? R_c has a frequency of 19, while in literary text O normally occurs 16 times and in telegraphic text 17 times. Let us try out the values for S, D, and O.

```
                                   5                                        1Ø
FDUHV  KRXOG  EHHAH  UFLVH  GWKDW  RQOBR  QHHGL  WLRQR  IDFRG  HRIDS
  RES   O D    EE E   R SE  D       O  O    EED     O O     OD  EO

                                  15                                        2Ø
DUWLF  XODUF  ODVVR  URQOB  RQHFL  SKHUN  HBLVE  HLQJX  VHGZL  WKLQD
 R       R     SSO   RO     O E     ER    E S    E      SED

                                  25                                        3Ø
XQLWD  WRQHW  LPHZK  HQDFR  GHLVU  HSODF  HGEBD  QHZFR  GHRUD  QHZHG
       O E    E     E  O    DE SR  E      ED      E O     DEOR    E ED

                                  35                                        4Ø
LWLRQ  WKHUH  SODFH  GFRGH  ZLOOE  HGHVW  URBHG  EBEXU  QLQJX  QOHVV
   O     ERE     E   D ODE          EDES   RO ED      R              ESS

RWKHU  ZLVHR  UGHUH  G.
O ER    SEO  RDERE  D
```

Thus far our deductions seem to be correct. In the cipher text FRGH, a 4-letter group occurs four times in the following 5-letter sequences of the message: 9-1Ø, 24-25, 28-29, and 34. We have tentatively RGH as ODE and the word CODE is immediately suggested. If this is right, another plain text letter is identified

and F_c equals C_p. Can we confirm it? Let us examine the 3rd, 4th, and 5th groups of five letters in the cryptogram.

They are EHHAH UFLVH GWKDW. Now take the sequence HAH UFLVH G. If A_c equals X_p,
EE E RC SE D E E RC SE D

and L_c equals I_p, we have HAH UFLVH G. A_c equalling X_p is certainly very good,
EXE RCISE D

because A occurs only once which is normal for X. The assignment of L_c for I_p appears to fit in. We will insert these values.

FDUHV KRXOG EHHAH UFLVH GWKDW RQOBR QHHGL WLRQR IDFRG HRIDS
C RES O D EEXE RCISE D O O EEDI IO O COD EO

DUWLF XQDUF ODVVR URQOB RQHFL SKHUN HBLVE HLQJX VHGZL WKLQD
R IC RC SSO RO O ECI ER E IS EI SED I I

XQLWD WRQHW LPHZK HQDFR GHLVU HSODF HGEBD QHZFR GHRUD QHZHD
I O E I E E CO DEISR E C ED E CO DEOR E E

LWLRQ WKHUH SODFH GFRGH ZLOOE HGHVW URBHG EBEXU QLQJX QOHVV
I IO ERE CE DCODE I EDES RO ED R I . ESS

RWKHU ZLVHR UGHUE G.
O ER ISEO RDERE D

A sequence of nine letters, just preceding the final group of seven tentatively translated as ORDERED, is RWKHU ZLVH. If we fill in the values- T, H, and W in
O ER ISE

place of W, K, and Z, we have the word OTHERWISE. This will give us three new letters of the clear text. In the fifth 5-letter sequence is WKDW . This indi-
TH T

cates that D_c equals A_p and our word is THAT. The 15th and 16th five-letter sequences of the cryptogram are RQHFL SKHUN and this suggests that S_c equals P_p,
O ECI HER

yielding the plain text CIPHER. Another intriguing sequence, V KRXOG EH , the 5th
S HO D E

to 12th letters of the message, offers as plain text SHOULD BE. This will give three additional values and we will now substitute T, H, W, A, P, U, L, and B in place of W, K, Z, D, S, X, O, and E, respectively.

All of the values assigned thus far tie in very well and confirm the assumptions we have made. Our deciphering alphabet, with the 16 clear text values is:

Cipher Component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Plain Component - X A B C D E H I L O P R S T U W

Only a brief survey of this deciphering alphabet is needed to see that the normal alphabet has simply been shifted three places and the sequences already established are sufficient to enable us to fill in the remaining ten letters of the plain component in their proper locations. The two alphabets for enciphering and deciphering are:

Enciphering Alphabet

Plain Component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Cipher Component - D E F G H I J K L M N O P Q R S T U V W X Y Z A B C

Deciphering Alphabet

Cipher Component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Plain Component - X Y Z A B C D E F G H I J K L M N O P Q R S T U V W

This is the system usually known as Julius Caesar's method.

The plain text is: CARE SHOULD BE EXERCISED THAT ONLY ONE EDITION OF A CODE OF A PARTICULAR CLASS OR ONLY ONE CIPHER KEY IS BEING USED WITHIN A UNIT AT ONE TIME. WHEN A CODE IS REPLACED BY A NEW CODE OR A NEW EDITION THE REPLACED CODE WILL BE DESTROYED BY BURNING UNLESS OTHERWISE ORDERED.

The above solution was arrived at by frequency analysis and the fitting of letter to letter.

The Probable Word

Let us again write the cryptogram, underlining repetitions of the longer sequences, also possible pattern words.

FDUHV KRXOG EHHAH UFLVH GWKDW RQOBR QHHGL WLRQR IDFRG HRIDS

DUWLF XODUF ODVVR URQOB RQHFL SKHUN HBLVE HLQJX VHGZL WKLQD

XQLWD WRQHW LPHZK HQDFR GHLVU HSODF HGEBD QHZFR GHRUD QHZHG

LWLRQ WKHUH SODFH GFRGH ZLOOE 'GHVW URBHG EBEXU QLQJX QOHVV

RWKHU ZLVHR UGHUH G.

With no collateral information, it is necessary to seek repeated and fairly long pattern words (idiomorphs), also repeated trigraphs. Available lists of such words must be checked for possibilities. See Chapter XIII.

Solution by Comparison of Frequency Graphs

In this problem the cryptogram could have been solved by sliding the two frequency graphs (for normal telegraph text and for the cipher text) against one another until they coincided, with D of the graph for the cryptogram opposite A of the graph for the normal.

NORMAL TEXT (Telegraph)	CIPHER TEXT
	A - 1
	B - 111111
	C -
A - 1111111111111111	D - 111111111111
B - 111	E - 11111
C - 111111	F - 1111111111
D - 11111111	G - 1111111111111
E - 11111111111111111111111111	H - 1111111111111111111111111111111
F - 1111	I - 11
G - 1111	J - 11
H - 11111111	K - 111111
I - 111111111111	L - 1111111111111111
J - 1	M -
K - 11	N - 1
L - 11111111	O - 111111111
M - 111111	P - 1
N - 1111111111111111	Q - 11111111111111
O - 11111111111111111	R - 111111111111111111
P - 11111	S - 1111
Q - 1	T -
R - 111111111111111	U - 111111111111
S - 11111111111111	V - 1111111111
T - 11111111111111	W - 111111111111
U - 111111	X - 111111
V - 111	Y -
W - 111	Z - 111111
X - 1	A - 1
Y - 1111	B - 111111
Z -	C -

The general coincidence of the two graphs is close enough to warrant the assumption that the cipher alphabet has simply been moved ahead 3 places.

Solution by Completing the Plain Component

When the frequency graph of a cryptogram leads us to believe that a direct standard alphabet, such as Caesar's in this case, has been used, we can save a great deal of time by completing the plain component, or "Running down the alphabet." Let us take the first 25 letters of our cryptogram and apply the process to that portion of the message we have solved.

	FDUHV	KRXOG	EHHAH	UFLVH	GWKDW
1	GEVIW	LSYPH	FIIBI	VGMWI	HXLEX
2	HFWJX	MTZQI	GJJCJ	WHNXJ	IYMFY
3	IGXKY	NUARJ	HKKDK	XIOYK	JZNGZ
4	JHYLZ	OVBSK	ILLEL	YJPZL	KAOHA
5	KIZMA	PWCTL	JMMFM	ZKQAM	LBPIB
6	LJANB	QXDUM	KNNGN	ALRBN	MCQJC
7	MKBOC	RYEVN	LOOHO	BMSCO	NDRKD
8	NLCPD	SZFWO	MPPIP	CNTDP	OESLE
9	OMDQE	TAGXP	NQQJQ	DOUEQ	PFTMF
1ø	PNERF	UBHYQ	ORRKR	EPVFR	QGUNG
11	QOFSQ	VCIZR	PSSLS	FQWGS	RHVOH
12	RPGTH	WDJAS	QTTMT	GRXHT	SIWPI
13	SQHUI	XEKBT	RUUNU	HSYIU	TJXQJ
14	TRIVJ	YFLCU	SVVOV	ITZJV	UKYRK
15	USJWK	ZGMDV	TWWPW	JUAKW	VLZSL
16	VTKXL	AHNEW	UXXQX	KVBLX	WMATM
17	WULYM	BIOFX	VYYRY	LWCMY	XNBUN
18	XVMZN	CJPGY	WZZSZ	MXDNZ	YOCVO
19	YWNAO	DKQHZ	XAATA	NYEOA	ZPDWP
2ø	ZXOBP	ELRIA	YBBUB	OZFPB	AQEXQ
21	AYPCQ	FMSJB	ZCCVC	PAGQC	BRFYR
22	BZQDR	GNTKC	ADDWD	QBHRD	CSGZS
23	CARES	HOULD	BEEXE	RCISE	DTHAT
24	DBSFT	IPVME	CFFYF	SDJTF	EUIBU
25	ECTGU	JQWNF	DGGZG	TEKUG	FVJCV

The 23rd line is the only one making complete sense. It is not necessary to run down the entire message, because the first 25 letters of the text yield us 15 plain text equivalents and the remaining 7 may be readily filled in.

The method of "running down the alphabet" may also be applied when it is suspected that a reversed standard alphabet has been used. In such a case there is one preliminary step required before the plain component is completed. Since a reversed sequence was used to produce the cryptogram, we must first convert the cryptogram into a standard sequence by means of the following alphabets:

Reversed Sequence - Z Y X W V U T S R Q P O N M L K J I H G F E D C B A ,

Standard Sequence - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z .

The letters of the cryptogram are found in the reversed sequence and are replaced by the letters immediately beneath them in the standard sequence. Suppose that the following is the cryptogram:

	O Y A Z N	C D O M F	O G T W O	G U C J K.	Converted,
it reads:					
	L B Z A M	X W L N U	L T G D L	T F X Q R	and is then
run down:					
1.	M C A B N	Y X M O V	M U H E M	U G Y R S	
2.	N D B C O	Z Y N P W	N V I F N	V H Z S T	
3.	O E C D P	A Z O Q X	O W J G O	W I A T U	
4.	P F D E Q	B A P R Y	P X K H P	X J B U V	
5.	Q G E F R	C B Q S Z	Q Y L I Q	Y K C V W	
6.	R H F G S	D C R T A	R Z M J R	Z L D W X	
7.	S I G H T	E D S U B	S A N K S	A M E X Y	
8.	T J H I U	F E T V C	T B O L T	B N F Y Z.	

It is unnecessary to proceed beyond this point, the seventh line showing what is clearly our message, "SIGHTED SUB SANK SAME." The X and Y at the end of the line are nulls to complete the 5-letter group for telegraph transmission.

It should be noted that we did not know with what letter the original reversed sequence commenced. However, all we had to do was to set Z to A in making our conversion. In this cryptogram the reversed alphabet began with the letter G.

Chapter XI

MONOALPHABETIC CIPHERS USING SYMBOLS AND NUMERALS

A cipher alphabet composed of symbols is sometimes used in monoalphabetic substitution, but because symbols cannot be transmitted by telegraphic means their employment is usually limited to communications that are to be sent through the mails, by messenger, or by similar channels.

Below is a sample of a cipher alphabet using the characters on a typewriter keyboard.

Enciphering Alphabet

Plain Component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Cipher Component - - 2 / : . 3 " ? 8 , # 4 $ 7 ¢ (6 ' 1 ; 5) % & 9 Ø

Using this alphabet, the message: "INFANTRY POINT RED COLUMN PASSED GREEN-FIELD JUNCTION AT SIX FORTY FIVE A. M. STOP REMAIN IN OBSERVATION REPORT WHEN MAIN BODY RED TROOPS SIGHTED," would yield the following cryptogram in normal word lengths:

8 7 3 - 7 ; ' 9 (¢ 8 7 ; ' . : / ¢ 4 5 $ 7

(- 1 1 . : " ' . . 7 3 8 . 4 : , 5 7 / ; 8 ¢ 7 - ;

1 8 & 3 ¢ ' ; 9 3 8) . - $ 1 ; ¢ (' . $ - 8 7

8 7 ¢ 2 1 . ') - ; 8 ¢ 7 ' . (¢ ' ; % ? . 7

$ - 8 7 2 ¢ : 9 ' . : ; ' ¢ ¢ (1 1 8 " ? ; . :

Decipherment of messages by the addressee is obviously a simple matter of replacing the symbols in the cryptogram by the letters of the plain text which they represent, according to the deciphering alphabet:

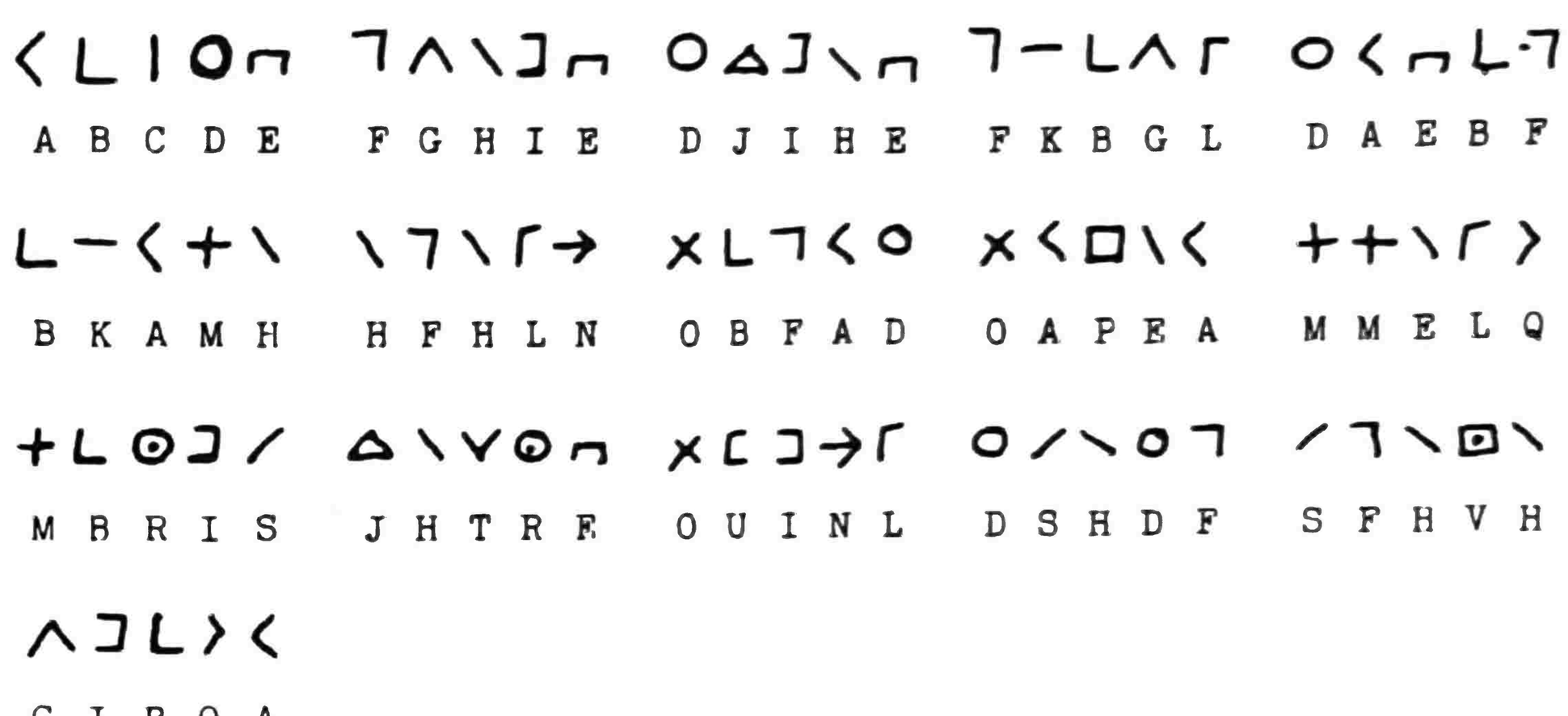
Cipher Component - 1 2 3 4 5 6 7 8 9 Ø . , : ; ' " - / ( ) ? # $ ¢ & %

Plain Component - S B F L U Q N I Y Z E J D T R G A C P V H K M O X W.

Cryptanalysis of ciphers employing symbols will be facilitated if letters are substituted for the various characters in the cryptogram.

Cryptanalysis of a Cipher Using Symbols

The following cryptogram, using symbols, has been intercepted and is to be solved. A rapid count of the number of characters in the message discloses that 22 different symbols have been employed. This indicates a probability of monoalphabetic substitution and to facilitate handling we will arbitrarily assign letters to each symbol and place them beneath the symbols in the cryptogram.



A B C D E F G H I E D J I H E F K B G L D A E B F

B K A M H H F H L N O B F A D O A P E A M M E L Q

M B R I S J H T R E O U I N L D S H D F S F H V H

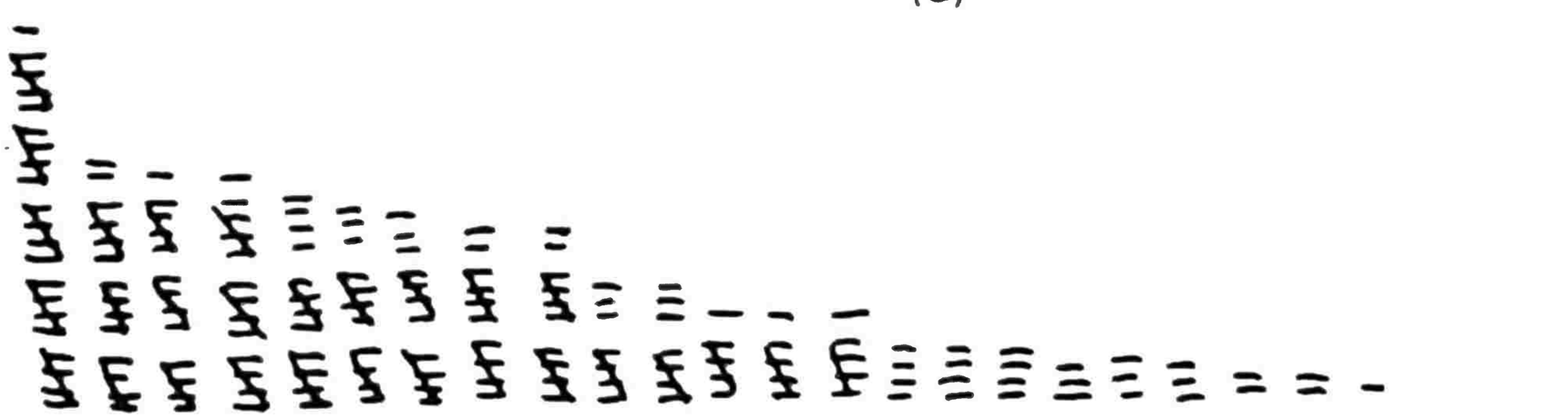
G I B Q A.

FIGURE 27

We will now draw up a single frequency chart with these substituted letters (Figure 28, (a)) below and compare it with the chart for normal literary frequency, (b) in Figure 28. In this instance we will arrange the letters in order of frequency of occurrence.

H A B E F D I L M G O R S J K N Q C P T U V

(a)



E T O A N I R S H D L U C M P F Y W G B V K J X Q Z

(b)

FIGURE 28

The next step is to substitute the letters of (b) in Figure 28 for those of (a), thus: H will be replaced by E, A by T, B by O, and so on.

				5	
A B C D E	F G H I E	D J I H E	F K B G L	D A E B F	Cipher
T O W I A	N D E R A	I M R E A	N P O D S	I T A O N	Plain
				1ø	
B K A M H	H F H L N	O B F A D	O A P E A	M M E L Q	Cipher
O P T H E	E N E S F	L O N T I	L T G A T	H H A S Y	Plain
				15	
M B R I S	J H T R E	O U I N L	D S H D F	S F H V H	Cipher
H O U R C	M E B U A	L V R F S	I C E I N	C N E K E	Plain

G I B Q A Cipher
D R O Y T. Plain

As no collateral information was supplied, we shall have to make use of the painstaking "trial and error" method and the use of the "probable word" along with it.

In the fifth group of five letters is the sequence ITAON. This is not normal in English, but if we transpose the A and the I we get ATION which is excellent. We therefore change the values so that D_c equals A_p, and E_c equals l_p. Another interesting sequence THE ENESF in groups 6 and 7, which may be THE ENEMY, in which case L_c and N_c equal M_p and Y_p, respectively. Again, in the 8th and 9th groups there is LONTI LT. With the transposition of A and I this would be LONTA LT, and if C is put in place of L it would yield the word CONTACT. These values will be substituted and the following will be our plain text as thus far decrypted:

TOWAI NDERI. A_REI NPODM ATION OPTHE ENEMY CONTA CTGIT HHIM_

HOUR_ _EBUI CVRYM A_EAN _NEKE DROYT.

The 3rd, 4th and 5th groups now show as - A_REI NPODM ATION. If we change the P and D to F and R in group 4, we will have the word INFORMATION, which looks good. In the second and third 5-letter groups, if we change D to R and R to L we will have NRELI A_LEI, and the word that stares us in the face is RELIABLE. Then in the 11th and 12th groups we will find HOUL_ BE_UI and in the 13th CVLYM. The two blanks in the 14th and 15th groups will probably take D, yielding MADE and

AND. The 11th, 12th, and 13th can easily be assumed to be HOULD BEQUI CKLYM, while the last letter in group 1Ø is clearly S. The rest of the solution follows without much difficulty and the clear text is:

"TO GAIN RELIABLE INFORMATION OF THE ENEMY CONTACT WITH HIM SHOULD BE QUICKLY MADE AND NEVER LOST."

Systems Employing Numerals

Cipher systems employing numerals are met with occasionally. Since there are only ten figures, the nine digits and zero, it is evident that each letter in the alphabet in the plain text must be represented in the cipher by at least two digits. The simplest form for an alphabet composed of numerals is one in which A equals 01, B is 02, C equals 03, and so on. The pairs of digits may be chosen systematically or at random. Two such alphabets are shown in Figure 29, below.

Systematic

Plain	Cipher
A	11
B	12
C	13
D	14
E	15
F	16
G	17
H	18
I	19
J	2Ø
K	21
L	22
M	23
N	24
O	25
P	26
Q	27
R	28
S	29
T	3Ø
U	31
V	32
W	33
X	34
Y	35
Z	36

(a)

Random

Plain	Cipher
A	37
B	65
C	24
D	18
E	11
F	Ø4
G	75
H	29
I	15
J	83
K	52
L	1Ø
M	44
N	16
O	27
P	56
Q	13
R	28
S	91
T	14
U	Ø5
V	39
W	78
X	49
Y	33
Z	Ø1

(b)

FIGURE 29

Using the plain text: "CRYPTOGRAPHIC MESSAGES SHOULD BE SHORT AND CONCISE AS LONG MESSAGES ARE EASIER FOR THE ENEMY TO SOLVE," and the systematically constructed alphabet shown in (a) of Figure 29 on page 71, the cryptogram in normal word lengths would be:

13 28 35 26 3Ø 25 17 28 11 26 18 19 13 23 15 29 29 11 17 15 29
29 18 25 31 22 14 12 15 29 18 25 28 3Ø 11 24 14
13 25 24 13 19 29 15 11 29 22 25 24 17 23 15 29 29 11 17 15 29
11 28 15 15 11 29 19 15 28 16 25 28 3Ø 18 15 15 24 15 23 35
3Ø 25 29 25 22 32 15.

In five-figure groups the cryptogram would appear thus:

13283 5263Ø 25172 81126 18191 32315 29291 11715 29291 82531
22141 21529 18252 83Ø11 24141 32524 13192 91511 29222 52417
23152 92911 17152 91128 15151 12919 15281 62528 3Ø181 51524
15233 53Ø25 29252 23215.

There a great many ways in which these numerical equivalents of letters may be constructed to form cipher alphabets. Figure 30 shows three of them. The numbers at the top are the column indicators and those at the sides of the square blocks are the row indicators.

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I-J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

(a)

	1	2	3	4	5
6	A	B	C	D	E
7	F	G	H	I-J	K
8	L	M	N	O	P
9	Q	R	S	T	U
Ø	V	W	X	Y	Z

(b)

	1	2	3	4	5	6	7
1	A	B	C	D	E	F	G
2	H	I	J	K	L	M	N
3	O	P	Q	R	S	T	U
4	V	W	X	Y	Z		

(c)

FIGURE 3Ø

In deriving the cipher alphabet, the figure of the column indicator is taken to be the first of the pair of digits representing the particular letter in that column, while the figure indicating the row in which the letter appears becomes the second digit of the pair. Thus, in (c) of Figure 3Ø on Page 72, the letter A is represented by 11 because it is in the first column and in the first row; 31 equals C, and 54 represents Z.

The French statesman, Mirabeau, devised a numerical cipher of this general type and it still bears his name. His system divides the alphabet into five groups of five letters each. One letter, I, J, U, or V, is usually omitted. A number is assigned to each group and to each letter in the group. Two varieties of the cipher are shown below.

1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
A B C D E	F G H I K	L M N O P	Q R S T U	V W X Y Z
1	2	3	4	5

(a)

1 4 3 2 5	1 4 3 2 5	1 4 3 2 5	1 4 3 2 5	1 4 3 2 5
A B C D E	F G H I J	K L M N O	P Q R S T	U W X Y Z
7	6	5	4	3

(b)

FIGURE 31

In enciphering, two digits represent each letter of the plain text, the number of the group being written as the first digit and the number assigned to the letter in that group as the second digit. The cryptogram may be written in pairs of digits, as fractions, or in 5-figure groups. Using the method in (a) if Figure 31, above, the plain text "SEND REINFORCEMENTS" is enciphered:

43 15 33 14 42 15 24 33 21 34 42 13 15 32 15 33 44 43, or

$\frac{4}{3}\ \frac{1}{5}\ \frac{3}{3}\ \frac{1}{4}$ $\frac{4}{2}\ \frac{1}{5}\ \frac{2}{4}\ \frac{3}{3}\ \frac{2}{1}\ \frac{3}{4}\ \frac{4}{2}\ \frac{1}{3}\ \frac{1}{5}\ \frac{3}{2}\ \frac{1}{5}\ \frac{3}{3}\ \frac{4}{4}\ \frac{4}{3}$ or

43153 31442 15243 32134 42131 53215 33444 3.

One more example of numerical cipher alphabets will be given on Page 74.

Another method of preparing an alphabet employs one digit for the column indicator, but two digits each for the row indicators, thus each letter in plain text is represented in the cryptogram by a number of three digits. In Figure 32 the letter J is omitted.

	1	2	3	4	5
16	A	B	C	D	E
17	F	G	H	I	K
18	L	M	N	O	P
19	Q	R	S	T	U
2Ø	V	W	X	Y	Z

FIGURE 32

The arrangement of the letters in the square blocks or the rectangular matrices is not limited to the examples given, but may be varied by the correspondents in any manner they desire. Mixed sequences, either random or systematic may be employed.

The method shown above produces a cryptogram that has three times as many characters as the clear text. The disadvantages of a system which requires a cipher text triple the length of the plain text preclude its use to any considerable extent. It means greater time and labor and a crowding of the channels of communication for transmission. The cryptographic security is low, being that of a single substitution.

Decipherment of all of these monoalphabetic systems employing numerical characters is simply the reverse of encipherment.

The same method is used in the cryptanalysis of numerical ciphers as described for cryptograms in which symbols form the characters of the secret text.

Chapter XII

MISCELLANEOUS MONOALPHABETIC CIPHERS,

POLY-LITERAL SYSTEMS

In the 16th century, Sir Francis Bacon invented his famous bi-literal cipher, the application of which to Elizabethan literature has caused many controversies concerning the authorship of many works - especially the plays and sonnets attributed to William Shakespeare.

Any two different things may be combined in groups of five to form 32 different combinations. The English alphabet during the Elizabethan Age contained 24 letters, I and J and U and V being equivalents. Bacon's alphabet is:

A - aaaaa	G - aabba	N - abbaa	T - baaba
B - aaaab	H - aabbb	O - abbab	UV - baabb
C - aaaba	IJ - abaaa	P - abbba	W - babaa
D - aaabb	K - abaab	Q - abbbb	X - babab
E - aabaa	L - ababa	R - baaaa	Y - babba
F - aabab	M - ababb	S - baaab	Z - babbb

Two forms of anything are used, one called the "a-form," the other the "b-form." Capital and small letters (upper and lower case), or Roman and Italic letters are two examples. The plain text or "open" message will have five times as many letters as the hidden communication. Using capitals to represent the "a-form" and small letters for the "b-form," the open text "REMAIN TILL I COME" is written in 5-letter groups

R E m A i	N t I l L	i C o m E
a a b a b	a b a b a	b a b b a

and the concealed message is F L Y.

Bacon's alphabet is systematically constructed and has a definite pattern. The first column contains sixteen a's, followed by eight b's, the latter stopping at the 24th letter, Z. The second column contains eight a's, then eight b's, and then eight a's again. Reading down the third column, we find four a's alternating with four b's. In the fourth column a and b occur alternately in pairs of the

same letter, while in the fifth column a and b are alternated.

Another bi-literal system is the one called the "Vowel" cipher. A square block or matrix containing 25 small square cells is constructed. One of the letters of the normal alphabet is placed in each cell, according to any arrangement the correspondents agree upon, one letter being omitted - usually I, J, U, or V. The five vowels are placed one above each column and one at the left of each row. These form the cipher equivalents for plain text letters, two vowels being used to encipher each letter of the clear. The vowel at the top of the column is taken first and that at the left of the row for the second of the pair. A in plain text is represented by AA, C by IA, and so on, with the arrangement in Figure 33.

	A	E	I	O	U
A	A	B	C	D	E
E	F	G	H	I	K
I	L	M	N	O	P
O	Q	R	S	T	U
U	V	W	X	Y	Z

FIGURE 33

The word, RADIO, would be enciphered: R A D I O

EO AA OA OE OI.

The variations of the vowel cipher, employing one and two key-words respectively, are shown below.

	W	H	I	T	E
F	A	B	C	D	E
L	F	G	H	I	K
O	L	M	N	O	P
U	Q	R	S	T	U
R	V	W	X	Y	Z

(a)

	B	R	O	W	N
B	Z	Y	X	W	V
R	Q	R	S	T	U
O	P	O	N	M	L
W	F	G	H	I	K
N	E	D	C	B	A

(b)

FIGURE 34

The arrangement of the column and row equivalents for the clear letters may be varied, of course, in any manner by agreement of the correspondents. These bi-literal systems may be recognized by the fact that the number of different characters in the cipher text is so small, five or ten.

Sir Francis Bacon's bi-literal cipher is bi-literal with regard to the number of different characters, while the other systems described are bi-literal with respect to the method of encipherment - two letters in cipher always representing a single letter in plain text.

A tri-literal system is one in which the cipher representation is in the ratio of three cipher letters for each letter of the clear. Here is one example:

Plain -	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	C	C	C	C	C	C	C	C
	A	A	A	B	C	B	C	B	C	B	A	C	A	B	B	C	C	B	C	A	B	A	C	B	C	A
	A	B	C	A	A	B	C	C	B	B	A	C	B	A	C	A	B	C	C	A	B	C	A	C	B	B

The word ENEMY would be cryptographed:

E	N	E	M	Y
ACA	BBA	ACA	BAB	CCB

In 5 - letter groups this would read ACABB AACAB ABCCB.

Monoalphabetic Substitution With Variants

In all of the cipher systems considered thus far only one cipher alphabet has been used and all of the cryptograms could be solved by the use of frequency data and analysis. The frequency of the characters of the cipher text has been identical with that of the clear. To make solution more difficult for the cryptanalyst many devices are resorted to, one of them being a method of having a choice of several different elements to encipher letters of the plain text, thus securing variable equivalents for the letters that occur more often than others and breaking up the usual frequency. This is known as "monoalphabetic substitution with variants."

There are a variety of ways of obtaining multiple representation and some of these will now be set forth.

The one hundred numbers ØØ to 99 are divided into sets of 25, one set for each letter of the alphabet, excluding J. Each group of 25 numbers occupies one of four columns, the arrangement in the column being governed by a 4-letter key-word placed above the columns. In a separate column at the left is the normal alphabet of the plain text. The cryptographer has a choice of any one of four equivalents for each letter in the clear, see Figure 35. Another system is shown in Figure 36.

	Key-word:			
Plain	S	A	L	E
A	Ø8	25	65	96
B	Ø9	26	66	97
C	1Ø	27	67	98
D	11	28	68	99
E	12	29	69	75
F	13	3Ø	7Ø	76
G	14	31	71	77
H	15	32	72	78
I	16	33	73	79
K	17	34	74	8Ø
L	18	35	5Ø	81
M	19	36	51	82
N	2Ø	37	52	83
O	21	38	53	84
P	22	39	54	85
Q	23	4Ø	55	86
R	24	41	56	87
S	ØØ	42	57	88
T	Ø1	43	58	89
U	Ø2	44	59	9Ø
V	Ø3	45	6Ø	91
W	Ø4	46	61	92
X	Ø5	47	62	93
Y	Ø6	48	63	94
Z	Ø7	49	64	95

FIGURE 35

	Key-word: B L U E							
B	ØØ	L	25	U	5Ø	E	75
C	Ø1	M	26	V	51	F	76
D	Ø2	N	27	W	52	G	77
E	Ø3	O	28	X	53	H	78
F	Ø4	P	29	Y	54	I	79
G	Ø5	Q	3Ø	Z	55	K	8Ø
H	Ø6	R	31	A	56	L	81
I	Ø7	S	32	B	57	M	82
K	Ø8	T	33	C	58	N	83
L	Ø9	U	34	D	59	O	84
M	1Ø	V	35	E	6Ø	P	85
N	11	W	36	F	61	Q	86
O	12	X	37	G	62	R	87
P	13	Y	38	H	63	S	88
Q	14	Z	39	I	64	T	89
R	15	A	4Ø	K	65	U	9Ø
S	16	B	41	L	66	V	91
T	17	C	42	M	67	W	92
U	18	D	43	N	68	X	93
V	19	E	44	O	69	Y	94
W	2Ø	F	45	P	7Ø	Z	95
X	21	G	46	Q	71	A	96
Y	22	H	47	R	72	B	97
Z	23	I	48	S	73	C	98
A	24	K	49	T	74	D	99

FIGURE 36

The letter S may be enciphered four different ways in each system. In Figure 35, the equivalents for S are ØØ, 42, 57, and 88. According to the alphabets in Figure 36, values for S are 16, 32, 73, and 88. With the key-word SALE, the four values for E are 12, 29, 69, and 75; and with the key-word BLUE - Ø3, 44, 6Ø, and 75. Two other methods of preparing variant alphabets are given on Page 79.

(a)

First Digit

(a)
Second Digit

	Ø	1	2	3	4	
5	A	B	C	D	E	Ø
6	F	G	H	I	K	1
7	L	M	N	O	P	2
8	Q	R	S	T	U	3
9	V	W	X	Y	Z	4
	5	6	7	8	9	

(b)
Second Digit

(b)

First Digit

FIGURE 37

S would be enciphered by <u>28 or 73</u>; E by <u>45 or 9Ø</u>.

(a) (b)

First Digit

(a)
Second Digit

	Ø	1	2	3	4	5	6	7	8	9	
Ø	U	N	I	T	E	D	E	C	L	A	Ø
1	D	S	A	O	F	R	T	I	O	N	1
2	M	R	C	B	G	F	P	B	G	H	2
3	H	K	L	P	Q	K	M	Q	S	U	3
4	V	W	X	Y	Z	V	W	X	Y	Z	4
5	C	O	N	S	T	U	N	I	O	A	5
6	I	U	A	L	Y	D	L	B	E	R	6
7	B	D	E	F	G	T	Y	C	F	G	7
8	H	K	M	P	Q	H	K	M	P	Q	8
9	R	V	W	X	Z	S	V	W	X	Z	9
	Ø	1	2	3	4	5	6	7	8	9	

(b)
Second Digit

(c)
Second Digit

(d)
Second Digit

(c) (d)

First Digit First Digit

FIGURE 38

Let us encipher a short message according to the method of Figure 36 on Page 78 with the key-word BLUE. The plain text is "REPORT FOR FURTHER ORDERS."

R E P O R T F O R F U R T H E R O R D E R S

15 Ø3 13 28 31 17 Ø4 12 72 45 18 87 33 Ø6 44 15 69 31 Ø2 6Ø 72 16

In the clear message there are seven R's, but in the cipher 15, 31 and 72, three of the values are repeated. There are three O's and three E's. The repeated digraph <u>OR</u> has three equivalents for O and two for R in the three occurrences and the cipher groups are 28 31, 12 72, and 69 31. Digraph ER occurs twice and its cipher values are 44 15 and 6Ø 72. There are only three repeated groups in the cryptogram as compared with the numerous repetitions in the plain text.

Ciphers of this type are sometimes called "homophonic" and the equivalents for the same letters "homophones." These systems are, of course, of somewhat greater security than single substitution methods, but they are susceptible of attack and can be solved without too much difficulty.

Using the key-word UNIFORMS and derived numerical key 8-4-2-1-5-6-3-7, shown on Page 10, we can construct another block with column and row indicators arranged according to this numerical key. The letters within the 25 cells of the block are placed in the normal order row by row (Figure 39).

			T	R	G	V	S
			J	X	U	A	K
F	P	B	A	B	C	D	E
D	M	L	F	G	H	I	K
Q	H	Z	L	M	N	O	P
I	W	O	Q	R	S	T	U
C	N	E	V	W	X	Y	Z

FIGURE 39

With this set-up, any plain text letter within the square matrix can be enciphered by six different pairs of digraphic (bi-literal) representations, but any digraph in the cipher text can have only one equivalent in the clear.

These "homophonic" systems form a link between monoalphabetic and poly-alphabetic substitution.

Chapter XIII

POLYALPHABETIC SYSTEMS

Polyalphabetic systems are those in which two or more cipher alphabets are used in converting the plain text into secret form. With three different cipher alphabets, the cryptographer enciphers the first letter of the clear from the 1st alphabet, the second letter from the 2nd alphabet, and the 3rd letter from the third alphabet. When he comes to the fourth letter of plain text, he uses the first alphabet again and similarly with the seventh, tenth and thirteenth letters. The second alphabet will be used for the second, fifth, and eighth letters, while the third alphabet will encipher the third, sixth, ninth and twelfth letters, and so on. This substitution is called periodic. It produces what we call "Cyclic Phenomena."

Let us use these three cipher alphabets and encipher a message, or rather - the first twenty-one letters of the text.

Plain Component - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Cipher
- 1st Alphabet - W N H P U Y S X Q Z R O T G L E M A J B F K D I V C
- 2nd Alphabet - R K V Q Z U S O P W B X T Y H I D A G M F C J L N E
- 3rd Alphabet - C Q A K W F T E N Y B P G I V D R M L X H U S O Z J

Plain	-	C	O	F		I	R	S		T	D	I		V	I	S		I	O	N		E	N	E		M	Y	F		Plain
Cipher	-	H	H	F		Q	A	L		B	Q	N		K	P	L		Q	H	I		U	Y	W		T	N	F		Cipher

The clear text is:

"C.O., FIRST DIVISION. ENEMY FORCE ESTIMATED TWO REGIMENTS INFANTRY WITH BATTALION OF LIGHT ARTILLERY NOW ADVANCING ON THOMASVILLE AT SEVEN A.M. HEAD OF COLUMN WAS TEN MILES EAST OF CROSS ROADS ONE FOUR TWO. REQUEST OUR ARTILLERY LAY DOWN INTERDICTION FIRE ON ALL APPROACHES TO ENEMY OBJECTIVE WHICH IS BELIEVED IS WEST OVER ENGINEERING FACTORY. JONES, BRIG. GEN."

Our cryptogram, with repeated groups underlined, is:

H H F Q A L B Q N K P L Q H I U Y W T N F L A A U Z L B P G 1

W M W P M S L A W S P G U Y X J P I Y R I B A Z D P X X K C 2

B M C O P V G H F O P T X M C A M N O X W A N I L J C P C C 3

G V N G S V G M E L T C J C N O X W W M L U C W G R G X Z C 4

P H F H H P F T I D R L B Z I T P P U G W W G X L U A A H L 5

J A V W Q L L Y W Y H H A M S L A W M F W J M V F A C A M N 6

O X W A N P W N K L J I Q Y X U A K Q V X Q H I Y P M U H I 7

W X P W I D A H C H O W J M V U Y W T N V N W W H M N K Z S 8

X P A X P L N Z P Q Z U U Q N J J W J M V K Z M U Y T Q Y W 9

U A N G S F W V X L A Z Z H I U C Q A P T S Z I. 1Ø

A survey of the cryptogram shows that there is a profusion of the rare consonants, J, K, Q, X, and Z. It is clearly a substitution cipher. Note the repetition of the 9-letter sequence, C AMNOX WAN in the third line and at the end of the sixth and beginning of the seventh lines; also the four-letter group NOXW in the fourth line.

How are we to determine the number of alphabets used, since it is suspected that the system employed is polyalphabetic? We have recourse to the process called "factoring." This consists of counting the intervals between repeated sequences of trigraphs and longer groups and obtaining the factors of the numbers of letters counted as intervening between repetitions of the same groups. In counting the total number of letters in the intervals, the letters of the group under consideration are included the first time that group occurs, but not the second. Below is the list of the repetitions, with the intervals and factors.

Lines	3-6,	CAMNOXWAN - CAMNOXWAN	-- 1Ø2 -	Factors:	3 × 2 × 17.
"	3-4,	NOXW - NOXW	-- 27 -	"	3 × 3 × 3.
"	4-6,	NOXW - NOXW	-- 72 -	"	3 × 3 × 2 × 2 × 2.
"	2-6,	MSLAW - MSLAW	-- 129 -	"	3 × 43.
"	1-8	UYWTN - UYWTN	-- 21Ø -	"	3 × 2 × 5 × 7.
"	6-8,	WJMV - WJMV	-- 51 -	"	3 × 17.
"	8-9,	WJMV - WJMV	-- 36 -	"	3 × 3 × 2 × 2.
"	1-7,	QHI - QHI	-- 189 -	"	3 × 3 × 3 × 7.
"	4-1Ø,	NGS - NGS	-- 180 -	"	3 × 3 × 2 × 2 × 5.

All other repeated trigraphs are parts of the longer sequences, so it is not necessary to list them separately.

The only factor common to all intervals is 3. Although there are other factors common to some of the repeated groups, the evidence favors the assumption that three cipher alphabets were employed.

The next step in the analysis of the message is the preparation of single frequency charts for the three alphabets.

	1st Alphabet		2nd Alphabet		3rd Alphabet	
A	11111 111	8	11111 11111	1Ø	111	3
B	11111	5				
C			111	3	11111 1111	9
D	11	2			1	1
E					1	1
F	11	2	1	1	11111	5
G	11111 11	7	111	3	111	3
H	1111	4	11111 11111 1	11	1	1
I			1	1	11111 11111 1	11
J	11111 11	7	111	3		
K	111	3	1	1	11	2
L	11111 1111	9			11111 111	8
M	1	1	11111 11111 111	13	11	2
N	11	2	11111	5	11111 1111	9
O	11111	5	1	1		
P	111	3	11111 11111 11	12	1111	4
Q	11111 11	7	111	3	11	2
R			111	3		
S	11	2	11	2	111	3
T	111	3	11	2	111	3
U	11111 11111 11	12	1	1	1	1
V			111	3	11111 11	7
W	11111 111	8	1	1	11111 11111 11111-1	16
X	11111	5	1111	4	11111 1	6
Y	111	3	11111 11	7		
Z	1	1	11111 111	8	11	2
	Totals	98		98		98

The appearances of the graphs for these alphabets show marked crests and troughs, as well as blanks. Our assumption of three alphabets seems to be correct.

Two letters stand out as probably E. They are U, with 12 occurrences, in the first alphabet; and W, with 16 appearances in the third. Although M is the most frequent letter in the second alphabet, there are three other letters that are closely pressing it for the lead - A, H, and P. Any one of the four might

represent E_p. Let us try out the 9-letter sequence and see what clues we can find. We will place the number of the alphabet above each letter. We will also list the other repeated groups.

	3 1 2 3 1 2 3 1 2 3
Cipher	C A M N O X W A N
Plain	E
Cipher	M S L A W
Plain	E
Cipher	U Y W T N
Plain	E E
Cipher	W J M V
Plain	E
Cipher	Q H I
Plain	
Cipher	N G S
Plain	

The third repetition, UYWTN, with U and W as E, separated by Y, looks possible for ENEMY. Y, T, and N are in alphabets 2 and 1, so we check their graphs to see if they are reasonable for N, M, and Y. The answer is yes, and we fill them in. A 9-letter word frequently encountered in military communications, with the ending E-Y, is ARTILLERY. Let us assume that this is the clear text for CAMNOXWAN and insert the plain equivalents.

	3 1 2 3 1 2 3 1 2 3
Cipher	C A M N O X W A N
Plain	A R T I L L E R Y
Cipher	M S L A W
Plain	T E
Cipher	U Y W T N
Plain	E N E M Y
Cipher	W J M V
Plain	E R
Cipher	Q H I
Plain	
Cipher	N G S
Plain	I

The trigraph NGS, with I--, suggests ING or ION. G in the first alphabet is good

for either N or O, but S in the second is low for N if the clear is ION. We therefore decide on NGS as ING. We will write the first three lines of our cryptogram and fill in the values thus far determined.

```
1 2 3 1 2   3 1 2 3 1   2 3 1 2 3   1 2 3 1 2   3 1 2 3 1   2 3 1 2 3
H H F Q A   L B Q N K   P L Q H I   U Y W T N   F L A A U   Z L B P G    1
                                    E N E M Y
W M W P M   S L A W S   P G U Y X   J P I Y R   I B A Z D   P X X K C    2
  T E   T       E       E N                                       A
B M C O P   V G H F O   P T X M C   A M N O X   W A N I L   J C P C C    3
  T A L     N     L       T A     R T I L L   E R Y       A     A
```

At the end of the second and beginning of the third lines we have the sequence

```
KC BMCOP VG.
 A  TAL  N
```

This immediately suggests BATTALION, and a check of the alphabets confirms the reasonableness of the assumption. The third line also has another group, between BATTALION and ARTILLERY: HFO PTXM. Why not try the sequence for
                                                                L T
OF LIGHT. It checks.

The rest of the analysis, based upon "probable words" and a careful check of each step and of the possible plain text values according to the frequency graphs of the three alphabets, proceeds without difficulty until the solution is reached.

Chapter XIV

SQUARE TABLES, DISCS AND SLIDE RULES.

THE VIGENÈRE CIPHER.

In 1586 Blaise de Vigenère, famous French diplomat, published his work, "Traicté des Chiffres," in which he described the cipher system that bears his name, although he, himself, did not claim to be the inventor. This well known method was employed during the American Revolution and in the Civil War as well as in later times. Modifications of it are still in use.

Referring to the figure below, the first horizontal alphabet, which is in capitals, is the alphabet of the plain text. A key-word is chosen and this key-word establishes the periodicity in the use of the cipher alphabets. The vertical column of capitals at the left is used for the letters of the key-word. If the key-word is BAG, the second, first and seventh alphabets will be used. Lower case letters denote the cipher alphabets.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
B	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a
C	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b
D	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c
E	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d
F	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e
G	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f
H	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g
I	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h
J	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i
K	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j
L	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k
M	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l
N	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m
O	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n
P	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
Q	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
R	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q
S	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
T	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
U	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t
V	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
W	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v
X	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w
Y	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x
Z	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y

FIGURE 40

Using the key-word BAG and the plain text COME BACK HERE, we write the letters of the key-word above the letters of the clear.

Key-word -	B A G B	A G B A	G B A G
Plain -	C O M E	B A C K	H E R E
Cipher -	D O S F	B G D K	N F R K.

The rule for enciphering is this: Find the letter of the plain text in the top horizontal line and the letter of the key-word in the vertical column at the left; the cipher letter will be found at the intersection of the vertical and horizontal lines drawn through the clear and key-letters respectively.

Decipherment is exactly the reverse of enciphering. The key-letter is located in the vertical column at the left, the cipher letter found in the same row as the key-letter; the plain text letter will be the one in the top horizontal row of capitals and in the same column as the cipher letter.

Porta's Table

Giovanni Battista della Porta, the famous Italian physician of the sixteenth century, devised the following table.

A B	a	b	c	d	e	f	g	h	i	j	k	l	m
	n	o	p	q	r	s	t	u	v	w	x	y	z
C D	a	b	c	d	e	f	g	h	i	j	k	l	m
	z	n	o	p	q	r	s	t	u	v	w	x	y
E F	a	b	c	d	e	f	g	h	i	j	k	l	m
	y	z	n	o	p	q	r	s	t	u	v	w	x
G H	a	b	c	d	e	f	g	h	i	j	k	l	m
	x	y	z	n	o	p	q	r	s	t	u	v	w
I J	a	b	c	d	e	f	g	h	i	j	k	l	m
	w	x	y	z	n	o	p	q	r	s	t	u	v
K L	a	b	c	d	e	f	g	h	i	j	k	l	m
	v	w	x	y	z	n	o	p	q	r	s	t	u
M N	a	b	c	d	e	f	g	h	i	j	k	l	m
	u	v	w	x	y	z	n	o	p	q	r	s	t
O P	a	b	c	d	e	f	g	h	i	j	k	l	m
	t	u	v	w	x	y	z	n	o	p	q	r	s
Q R	a	b	c	d	e	f	g	h	i	j	k	l	m
	s	t	u	v	w	x	y	z	n	o	p	q	r

S T	a	b	c	d	e	f	g	h	i	j	k	l	m
	r	s	t	u	v	w	x	y	z	n	o	p	q
U V	a	b	c	d	e	f	g	h	i	j	k	l	m
	q	r	s	t	u	v	w	x	y	z	n	o	p
W X	a	b	c	d	e	f	g	h	i	j	k	l	m
	p	q	r	s	t	u	v	w	x	y	z	n	o
Y Z	a	b	c	d	e	f	g	h	i	j	k	l	m
	o	p	q	r	s	t	u	v	w	x	y	z	n

FIGURE 41

The capital letters at the left are the letters of the key. The plain text letter is found in the two lines of lower case letters at the right of the key letter and the cipher letter will be the lower case letter above or below the clear letter as the case may be. There are really only 13 alphabets and the values are reciprocal for each set of 13 letters in each alphabet.

The Beaufort Table, devised by Admiral Beaufort, of the British Navy, is given in Figure 42, below.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A

FIGURE 42

Enciphering procedure is as follows: To encipher E with the key-letter M; find E in any one of the four alphabets that form the border, proceed up or down the column or to the right or left along the row until the key-letter, M, is reached, then go at right angles and the cipher letter will be found in the border alphabet as you leave the square. Decipherment is the reverse, of course, of the encipherment.

The signal disc formerly used in the United States Army is illustrated in Figure 43.

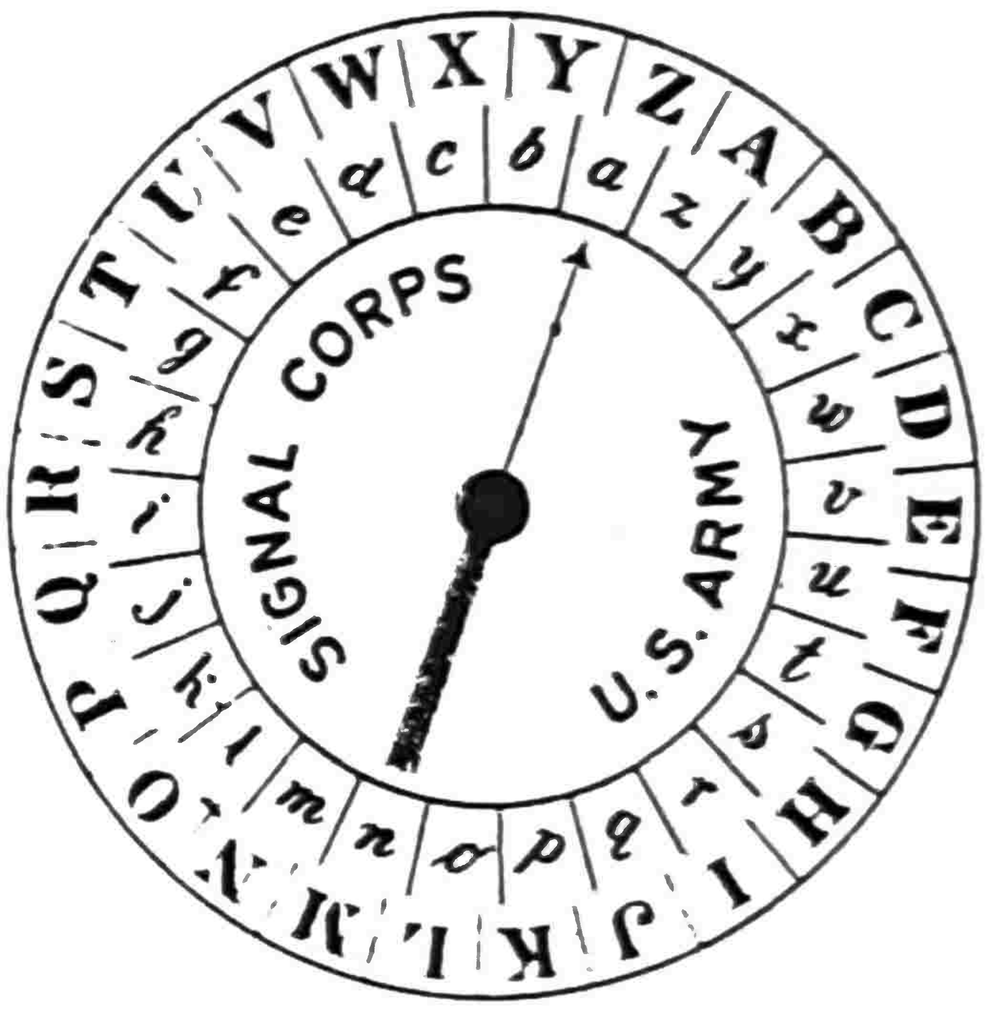


FIGURE 43

The key-letter is in the alphabet of the inner circle of the movable pa device, the plain text letter is in the outer alphabet. The key-letter opposite the A of the outer alphabet. Assuming that the plain text lett then the letter of the inner alphabet which comes opposite S becomes the letter. In this case it is M. Two strips of paper, on one of which the alphabet is written twice and on the other the reversed alphabet once, ma used as a substitute for the cipher disc.

Outer Alphabet - ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVWXYZ
ZYXWVUTSRQPONMLKJIH

Figure 44 shows the cipher disc and sliderule in the

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Z Y X W V U T S R Q P O N M L K J I H G F E D C B A
Y X W V U T S R Q P O N M L K J I H G F E D C B A Z
X W V U T S R Q P O N M L K J I H G F E D C B A Z Y
W V U T S R Q P O N M L K J I H G F E D C B A Z Y X
V U T S R Q P O N M L K J I H G F E D C B A Z Y X W
U T S R Q P O N M L K J I H G F E D C B A Z Y X W V
T S R Q P O N M L K J I H G F E D C B A Z Y X W V U
S R Q P O N M L K J I H G F E D C B A Z Y X W V U T
R Q P O N M L K J I H G F E D C B A Z Y X W V U T S
Q P O N M L K J I H G F E D C B A Z Y X W V U T S R
P O N M L K J I H G F E D C B A Z Y X W V U T S R Q
O N M L K J I H G F E D C B A Z Y X W V U T S R Q P
N M L K J I H G F E D C B A Z Y X W V U T S R Q P O
M L K J I H G F E D C B A Z Y X W V U T S R Q P O N
L K J I H G F E D C B A Z Y X W V U T S R Q P O N M
K J I H G F E D C B A Z Y X W V U T S R Q P O N M L
J I H G F E D C B A Z Y X W V U T S R Q P O N M L K
I H G F E D C B A Z Y X W V U T S R Q P O N M L K J
H G F E D C B A Z Y X W V U T S R Q P O N M L K J I
G F E D C B A Z Y X W V U T S R Q P O N M L K J I H
F E D C B A Z Y X W V U T S R Q P O N M L K J I H G
E D C B A Z Y X W V U T S R Q P O N M L K J I H G F
D C B A Z Y X W V U T S R Q P O N M L K J I H G F E
C B A Z Y X W V U T S R Q P O N M L K J I H G F E D
B A Z Y X W V U T S R Q P O N M L K J I H G F E D C
A Z Y X W V U T S R Q P O N M L K J I H G F E D C B
```

FIGURE 44

The alphabets of a Vigenere square may be classified as primary and secondary alphabets. The first lower case alphabet in Figure 40 is the primary alphabet and begins with the letter A. All of the other alphabets, commencing with the one beginning with the letter B, are secondary alphabets. This second alphabet is also known as the first generatrix. The alphabet beginning with C is the second generatrix.

The lower case alphabets in the Vigenere square need not be limited to the direct standard alphabets shown. Reversed standard and mixed, both systematic and random for the latter, may be used. A key-word mixed alphabet with the word LOGARITHMS would start off:

```
  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
L l o g a r i t h m s b c d e f j k n p q u v w x y z
O o g a r i t h m s b c d e f j k n p q u v w x y z l
G g a r i t h m s b c d e f j k n p q u v w x y z l o
A a r i t h m s b c d e f j k n p q u v w x y z l o g
R r i t h m s b c d e f j k n p q u v w x y z l o g a
```

```
I i t h m s b c d e f j k n p q u v w x y z l o g a r
T t h m s b c d e f j k n p q u v w x y z l o g a r i
H h m s b c d e f j k n p q u v w x y z l o g a r i t
M m s b c d e f j k n p q u v w x y z l o g a r i t h
S s b c d e f j k n p q u v w x y z l o g a r i t h m
B b c d e f j k n p q u v w x y z l o g a r i t h m s
C c d e f j k n p q u v w x y z l o g a r i t h m s b
D d e f j k n p q u v w x y z l o g a r i t h m s b c
E e f j k n p q u v w x y z l o g a r i t h m s b c d
F f j k n p q u v w x y z l o g a r i t h m s b c d e
J j k n p q u v w x y z l o g a r i t h m s b c d e f
K k n p q u v w x y z l o g a r i t h m s b c d e f j
N n p q u v w x y z l o g a r i t h m s b c d e f j k
P p q u v w x y z l o g a r i t h m s b c d e f j k n
Q q u v w x y z l o g a r i t h m s b c d e f j k n p
U u v w x y z l o g a r i t h m s b c d e f j k n p q
V v w x y z l o g a r i t h m s b c d e f j k n p q u
W w x y z l o g a r i t h m s b c d e f j k n p q u v
X x y z l o g a r i t h m s b c d e f j k n p q u v w
Y y z l o g a r i t h m s b c d e f j k n p q u v w x
Z z l o g a r i t h m s b c d e f j k n p q u v w x y
```

FIGURE 45

The following schematic diagrams illustrate graphically the processes of enciphering and deciphering with the Vigenère table and some of its modifications. Two alternative methods are shown for the Vigenère.

VIGENÈRE (1)

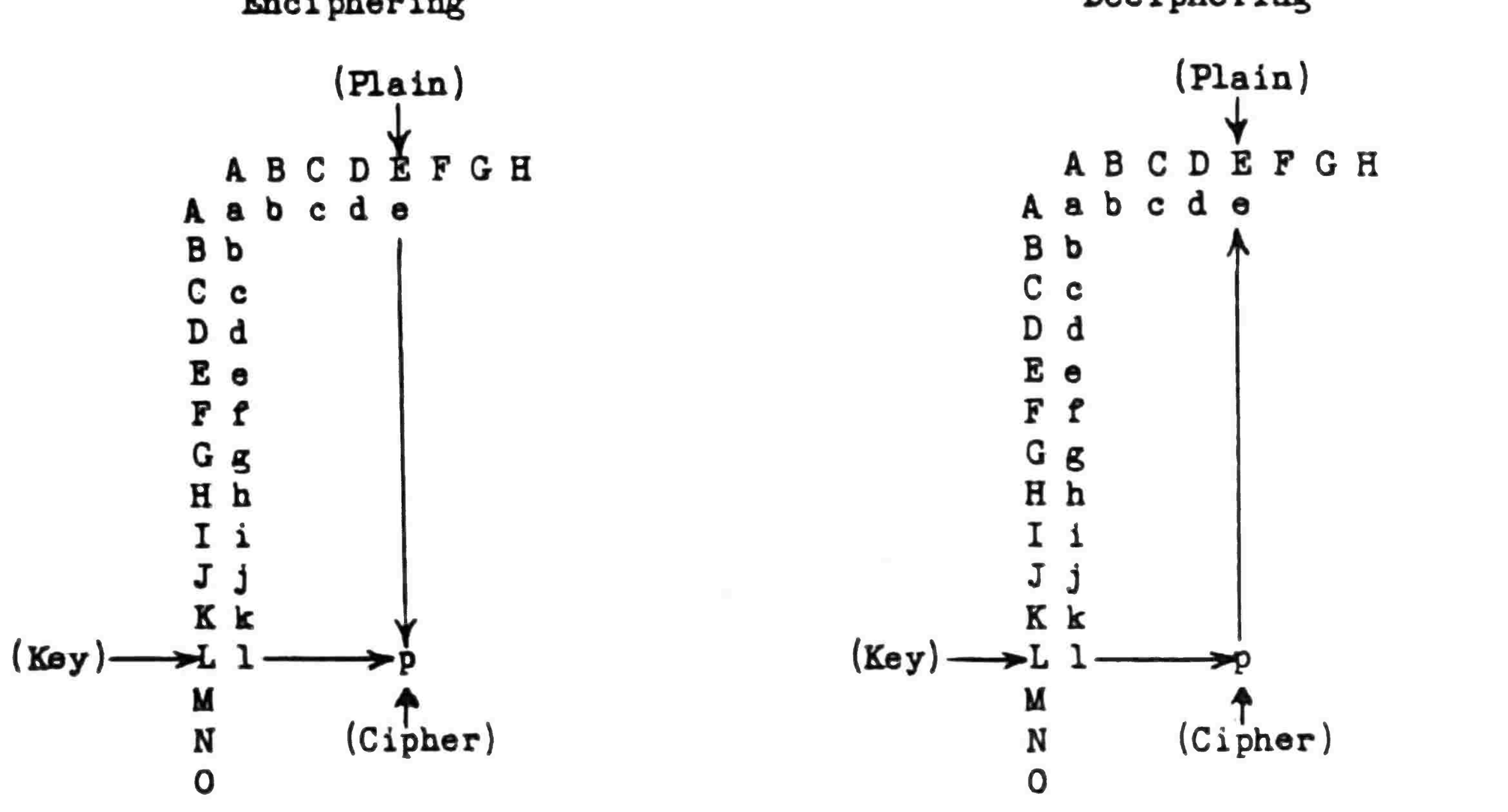


FIGURE 46

VIGENÈRE (2)

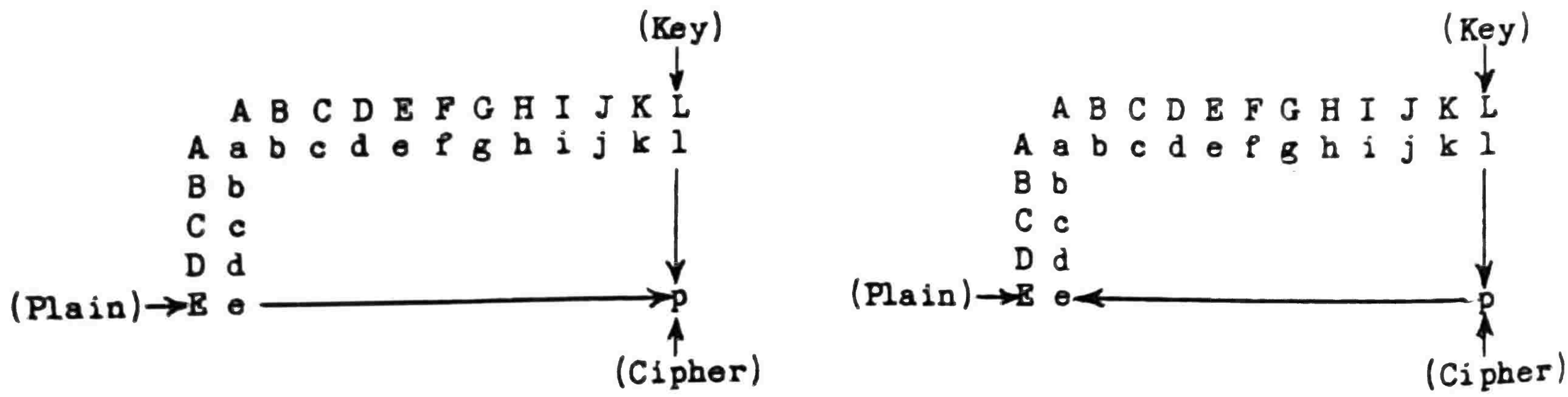

Enciphering
Deciphering
(Key)
A B C D E F G H I J K L
A a b c d e f g h i j k l
B b
C c
D d
(Plain) E e p
(Cipher)
(Key)
A B C D E F G H I J K L
A a b c d e f g h i j k l
B b
C c
D d
(Plain) E e p
(Cipher)

FIGURE 47

BEAUFORT (1)

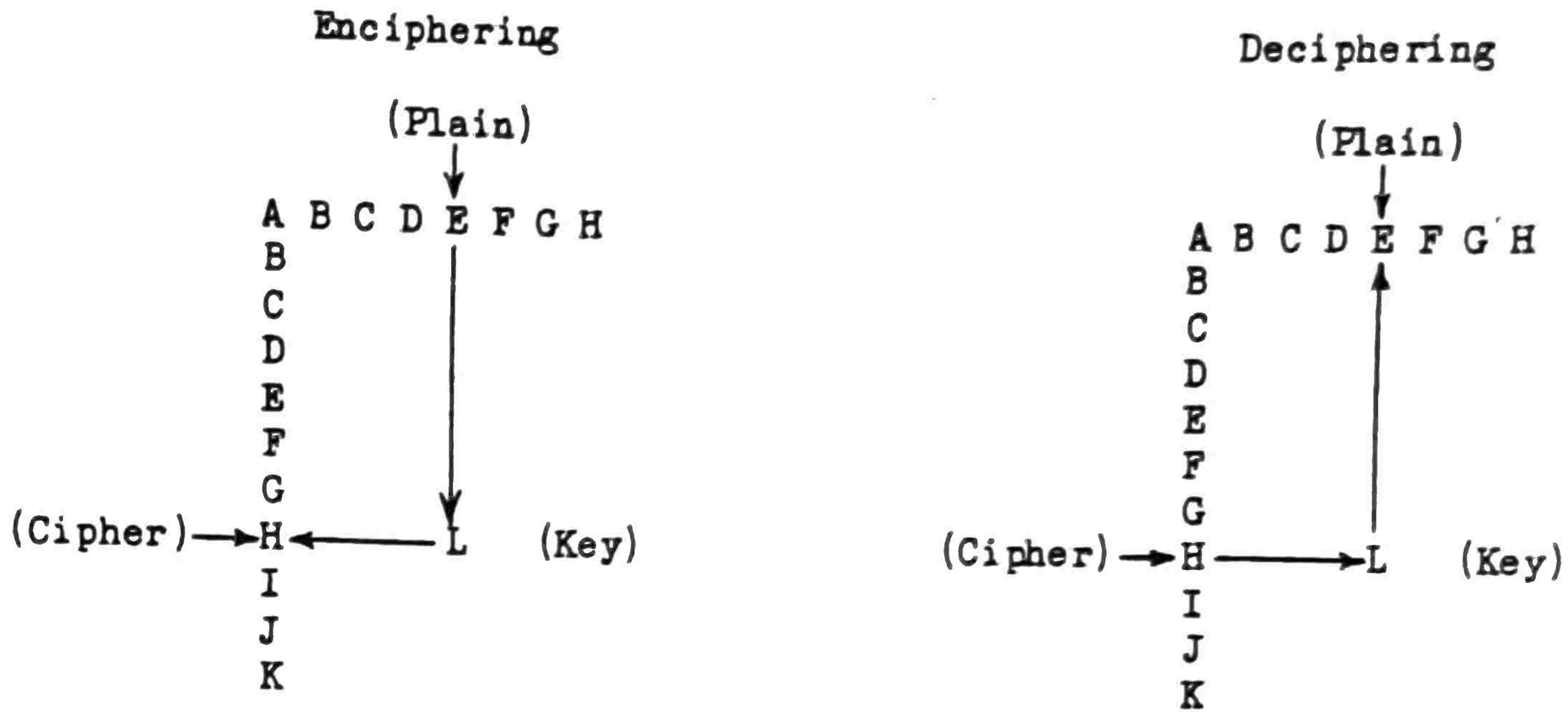


FIGURE 48

BEAUFORT (2)

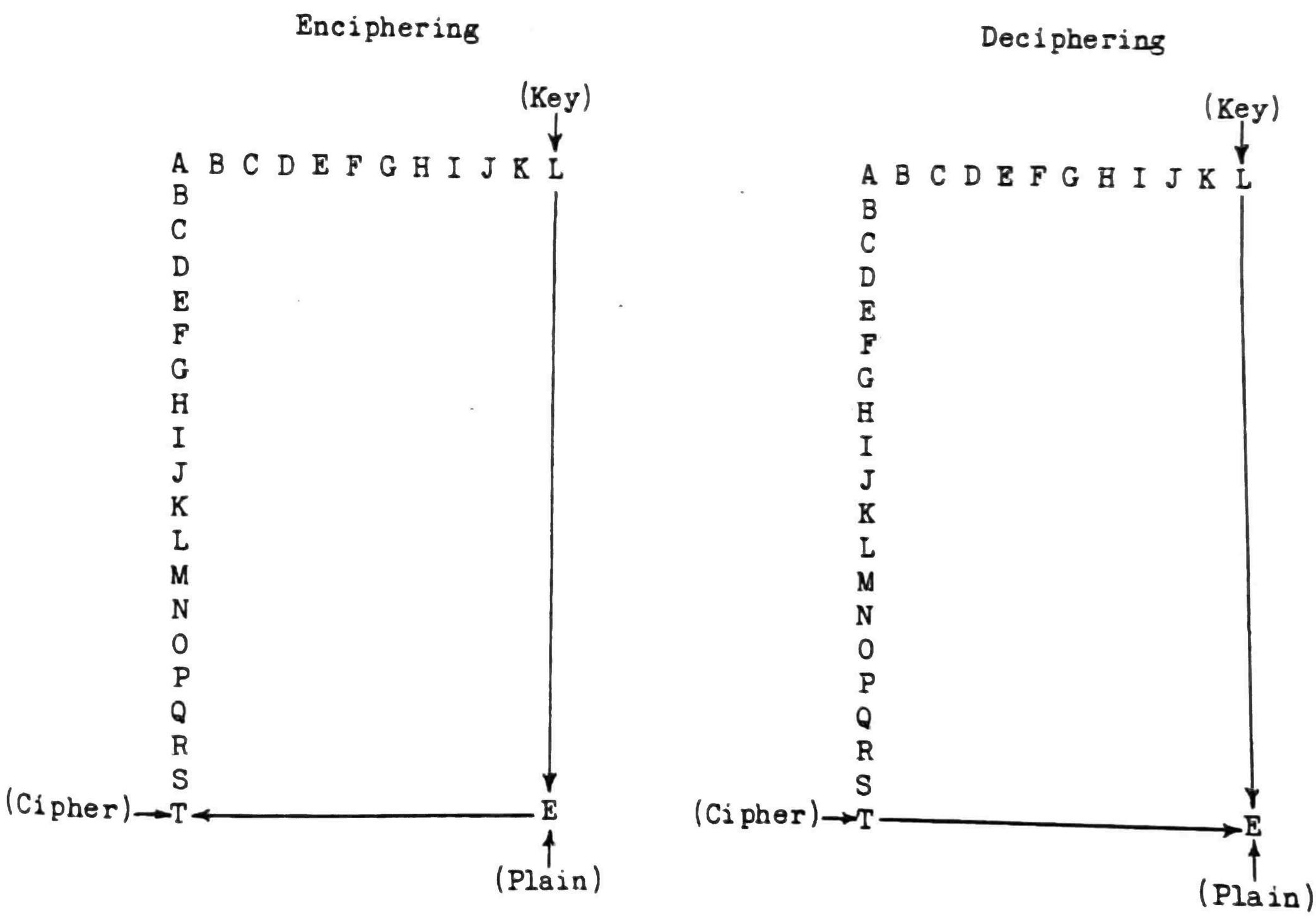


FIGURE 49

KEY-WORD MODIFIED VIGENÈRE

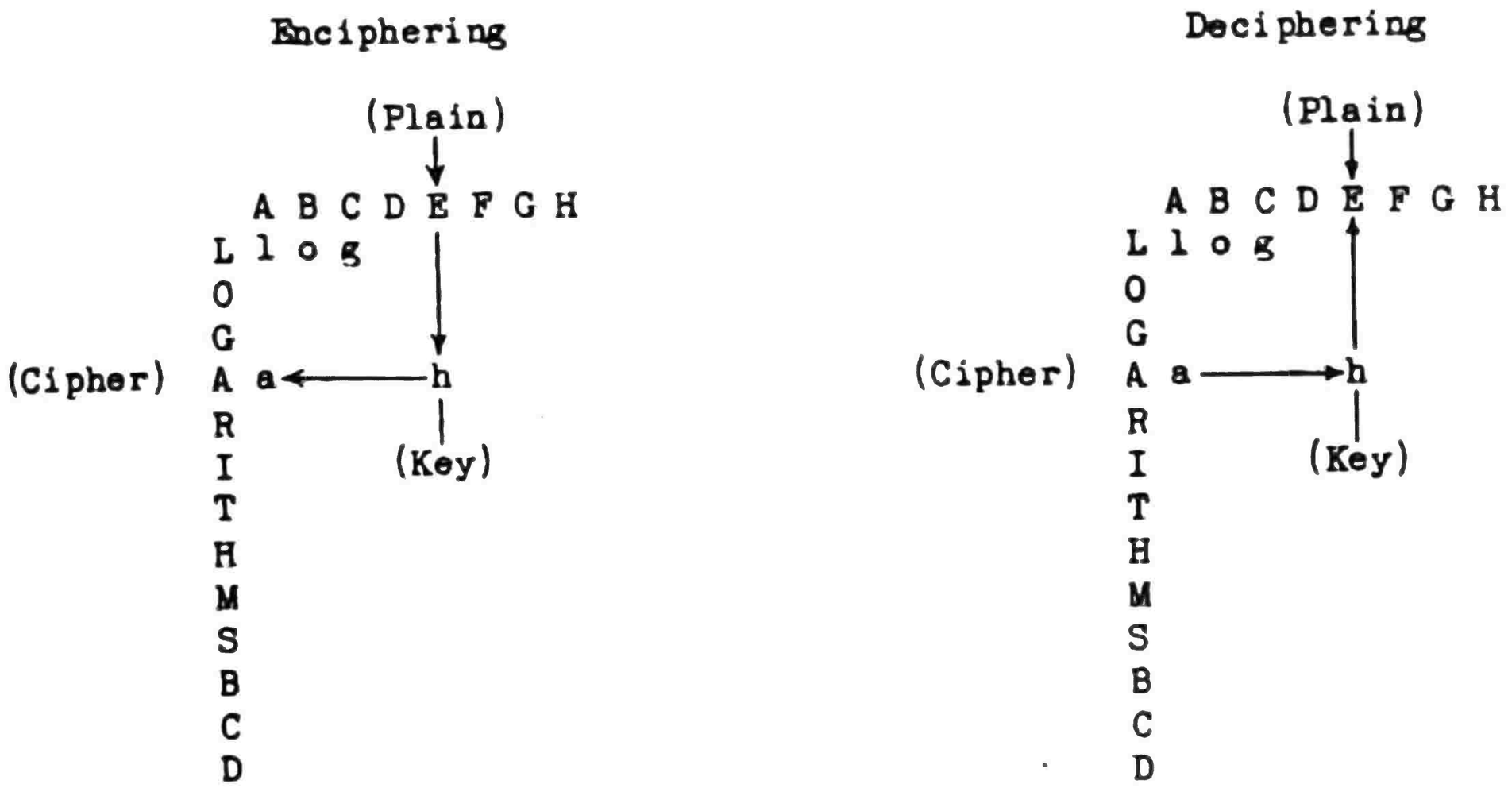


FIGURE 50

SIGNAL DISC

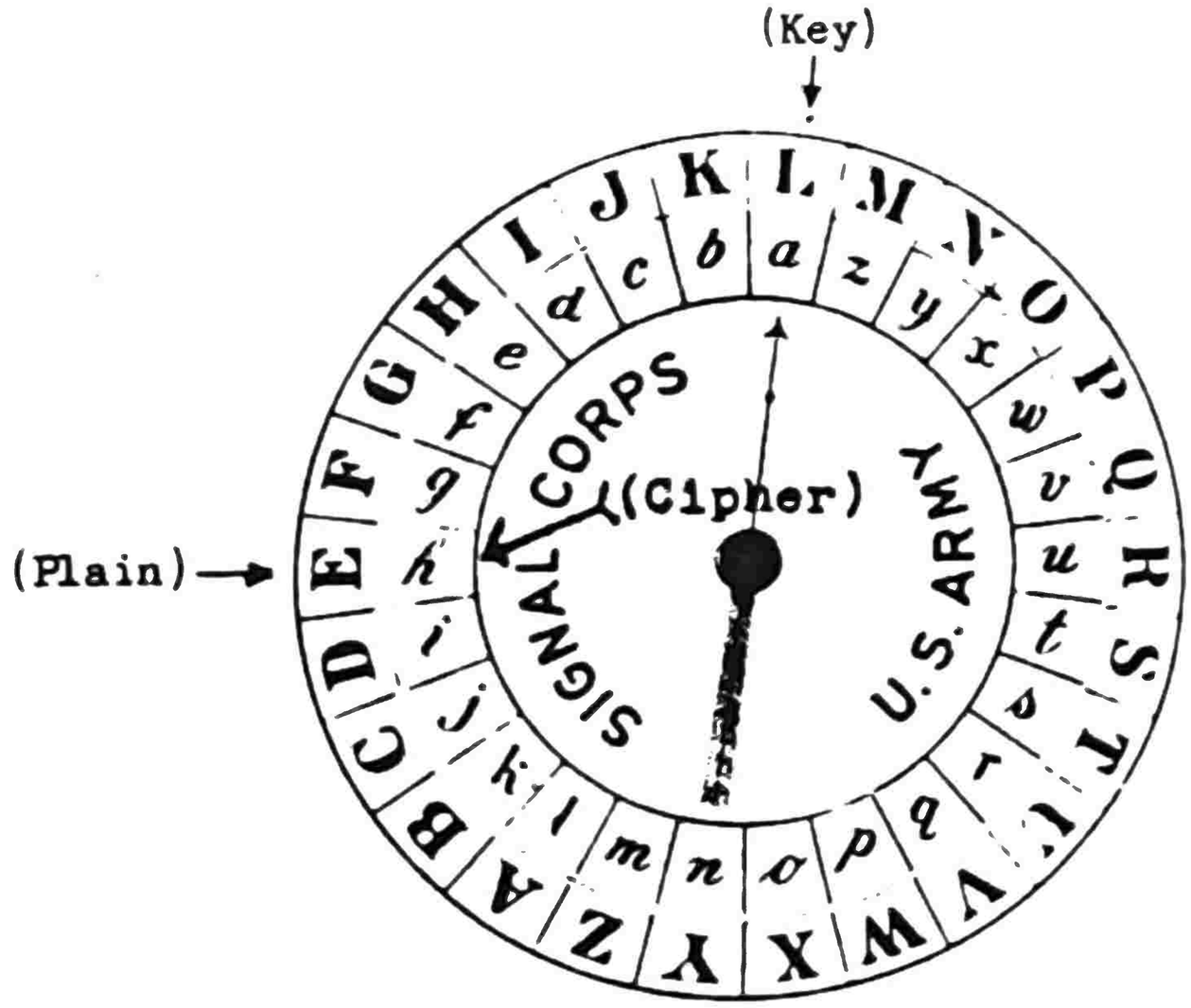


FIGURE 51

SLIDE-RULE EQUIVALENT

OF THE SIGNAL DISC.

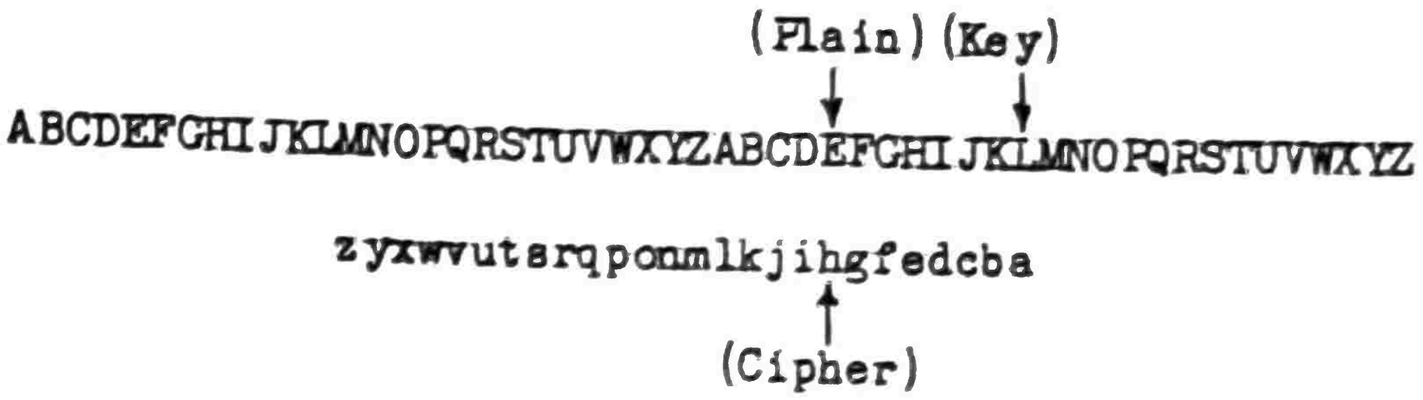


FIGURE 52

SQUARE TABLE EQUIVALENT

OF THE SIGNAL DISC.

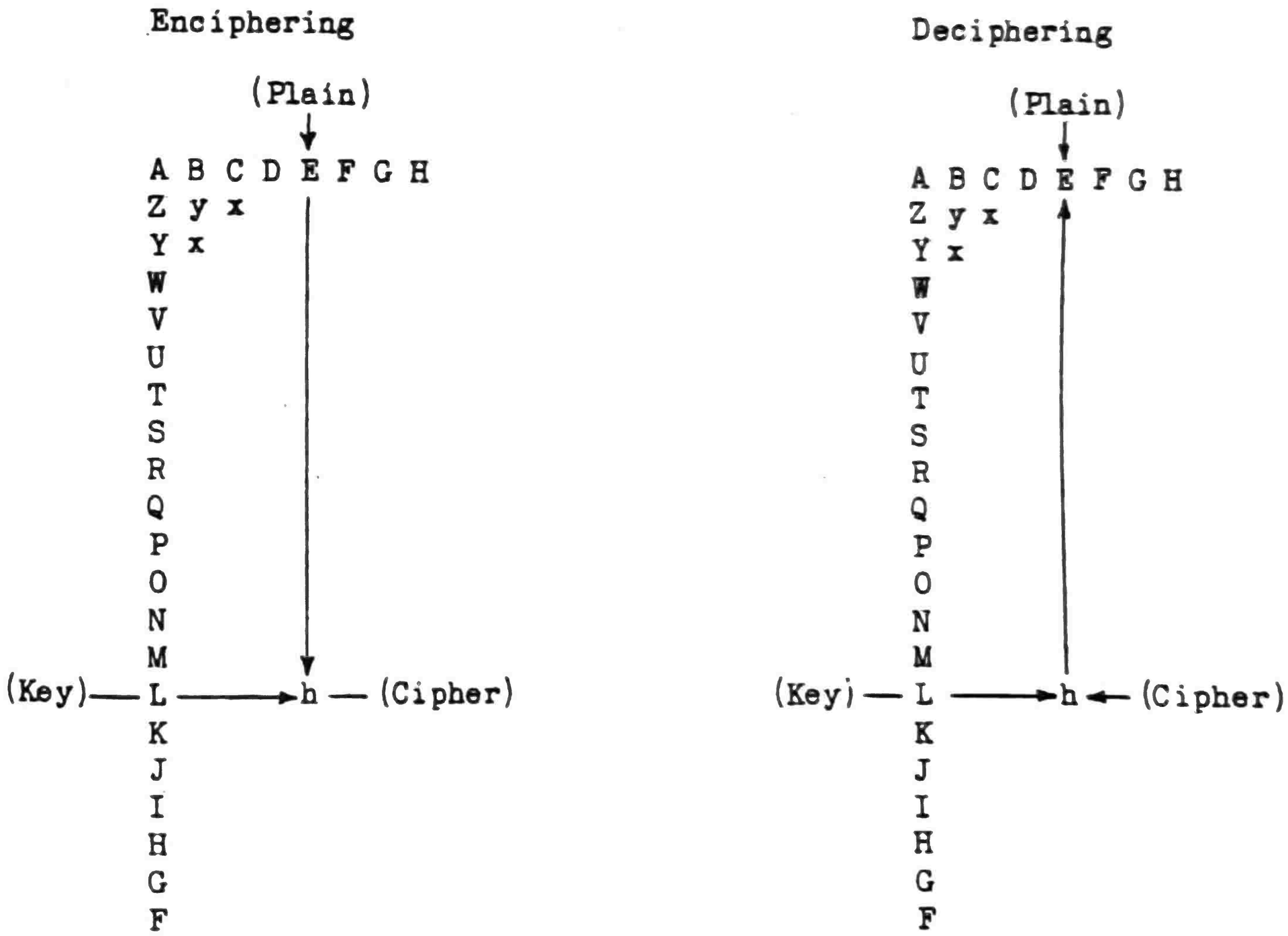


FIGURE 53

Chapter XV

DIGRAPHIC SYSTEMS.

THE PLAYFAIR.

The PLAYFAIR cipher is a digraphic substitution system used for a great many years as a field cipher in the British Army and employed for a time by both the American and British forces during the World War Number One. It was devised by the eminent English scientist, Sir Charles Wheatstone - inventor of the Wheatstone bridge - but derives its name from Sir Lionel Playfair who first sponsored its use.

Encipherment in a digraphic system is accomplished two letters at a time and each pair of letters in the clear text is represented by a pair of letters in the cipher text.

The Playfair cipher is based upon a square matrix, or block, composed of twenty-five small squares or cells in which the letters of the alphabet are written. There being but twenty-five cells, while the English alphabet contains twenty-six letters, I and J are usually placed in the same cell. The method of arranging the letters is determined by two factors, a key-word or key-phrase which determines the order of writing down the letters, and an agreement between the correspondents of the particular pattern or route to be followed. Key-words generally consist of ten letters in which there are no repetitions, e.g BANKRUPTCY, LOGARITHMS, PATRONIZED. The commonest practice is to begin in the upper left corner of the matrix and write the letters of the key-word one after another in the first two rows of cells. With the word BANKRUPTCY, the square is as follows:

B	A	N	K	R
U	P	T	C	Y
D	E	F	G	H
I-J	L	M	O	Q
X	V	W	X	Z

FIGURE 54

Below are four examples of other routes of inscribing the letters of the key-word:

1st and 3rd Rows	2d and 4th Rows	Columnar 1st and 2nd Cols.	Diagonally
B A N K R	D E F G H	B U D I S	B A K P D
D E F G H	B A N K R	A P E L V	N R T E I
U P T C Y	I L M O Q	N T F M W	U C D L Q
I L M O Q	U P T C Y	K C G O X	Y G M S W
S V W X Z	S V W X Z	R Y H Q Z	H O V X Z

FIGURE 55

Construction of a Playfair Square by Means of the Knight's Move in Chess

The Knight's move in the game of Chess is one square straight ahead and one square forward diagonally. This move may be used to construct a square matrix for the Playfair cipher. The starting point in the figure below is the center cell, Number 1. The other numbers, from 2 to 25, indicate the squares to which the Knight moves in sequences.

19	8	13	2	25
14	3	18	7	12
9	20	1	24	17
4	15	22	11	6
21	10	5	16	23

FIGURE 56

The letters of the chosen 10-letter key-word are written first in order in the cells numbered 1 to 10, followed by other letters of the alphabet not in the key-word in regular sequence in cells 11 to 25. Using the key-word BANKRUPTCY and omitting the letter J, the square will be:

O	T	F	A	Z
G	N	M	P	E
C	Q	B	X	L
K	H	V	D	U
S	Y	R	I	W

FIGURE 57

This is rather an effective way of mixing the letters of the alphabet. It may be used for the construction of systematically mixed alphabets for other methods of encipherment. From this matrix the letters may be transcribed according to any route agreed upon. Taking the rows from top to bottom, the mixed alphabet derived from the above square would be:

O T F A Z G N M P E C Q B X L K H V D U S Y R I W J.

J does not appear in the matrix, so it may be added at the end or at the beginning, or be placed in any position among the letters, as the correspondents may agree.

Rules for Enciphering

1. Divide the plain text message into groups of two letters. If the message contains an odd number of letters, such as 43, add another letter such as X or Z to complete the final digraph.
2. If doubled letters like EE or MM occur, separate them by inserting an X or Z between them. This cipher does not permit encipherment of a pair of like letters.
3. Each pair of letters in the square must be either (a) in the same horizontal row or (b) in the same vertical column, or (c) in the opposite corners of a rectangle formed by the small cells of the matrix.
4. When a pair of letters occurs in the same horizontal row, replace them in the cipher text by the letters to their right respectively. If one of the letters is at the extreme right of the row, the letter at the extreme left of the row is substituted. The rows and columns of letters are considered as forming rings or circles of letters.
5. When the pair of letters occurs in the same vertical column, the letters immediately below them are substituted. If one of the letters is at the bottom of the column, the letter at the top of that column replaces it.
6. When the letters of the digraph occur in the opposite corners of a rectangle, each of the letters is replaced by the letter in the other corner of the rectangle and in the same horizontal row.

Example of Encipherment

Using the matrix in Figure 46, with the key-word BANKRUPTCY written in the first and second rows, let us encipher the message,

"SEND REINFORCEMENTS AND AMMUNITION IMMEDIATELY."

We first divide the plain text into two-letter groups.

SE ND RE IN FO RC EM EN TS AN DA MX MU NI TI ON IM ME DI AT EL YX.

Note that the letter X had been inserted between the two M's of the word AMMUNITION and that another X had been added at the end to complete the digraph. Both of the X's are underlined.

The first digraph is SE and the two letters are at opposite corners of a rectangle. The substitutions are V from the other corner and in the same row as S, and D which is in the same row as E, giving us the combination VD. The next pair ND is likewise found at opposite corners of a rectangle and the substitution for these two letters is BF. The third digraph is RE, also in a rectangle and similarly enciphered by AH. The succeeding pairs of letters through the ninth, TS, follow the same procedure. The tenth digraph AN has its letters in the same horizontal row. Here the letters to the right of A and N are N and K, respectively, and the cipher group is NK. The complete encipherment is:

Plain - SE ND RE IN FO RC EM EN TS AN DA MX MU NI TI ON IM ME DI AT EL YX.

Cipher - VD BF AH MB GM KY FL FA UW NK EB OW IT BM UM MK LO LF IS NO LV CZ.

Note that the fourth from last digraph, DI, has the letters which are in the same vertical column, therefore D is enciphered by I just below it and I by S. Similarly LV is substituted for EL as the next to last pair. Because there are twenty-two digraphs totalling 44 letters in the cipher text, when this is arranged in five-letter groups for telegraphic transmission, we must add another null at the end to complete the final group and make it consist of five letters. Our cryptogram as it goes over the wire is:

VDBFA HMBGM KYFLF AUWNK EBOWI TBMUM MKLOL FISNO LVCZX.

Decipherment

Deciphering of the message by the recipient is simply the reverse of the procedure used in enciphering. Since the addressee knows that the system of enciphering is the PLAYFAIR, he counts the letters in the cryptogram and notes that the total is 45. He therefore drops the final letter X as a null and then divides the 44 letters of the cipher into two-letter groups. In decipherment the following rules apply:

1. To avoid errors, all nulls are written down as deciphered and when decipherment is completed they are crossed out as soon as the addressee decides he has the message properly converted into plain text.
2. When letters of a digraph are in opposite corners of a rectangle their clear text equivalents are in the other corners and in the same rows respectively, just as in enciphering.
3. If the letters of the pair are in the same horizontal row, then the plain language equivalents are just to their respective lefts, the reverse of enciphering, the one at the extreme right replacing the one at the extreme left.
4. When two letters are in the same vertical column, the equivalents are found just above each letter, except that when a letter is at the top of the column its equivalent in plain text will be at the bottom of the same column.

Peculiarities of the Playfair and Rules for Recognizing and Solving Cryptograms Enciphered by the Playfair System

1. It is a substitution system, therefore the rare consonants, J, K, Q, X, and Z, may be expected in the cryptogram with higher frequencies than normally occur. Digraphs containing them will recur in higher percentages.
2. There is always an even number of letters.
3. Only 25 different letters of the alphabet appear in the cryptogram, except nulls.
4. When the cipher text is broken up into digraphs, doubled letters, such as EE, MM, and SS, will not appear as digraphs.
5. Except in very short messages, digraphs will be repeated and their recurrences will generally follow the normal frequency.
6. In long messages whole words or phrases are likely to recur in repeated groups.
7. No single letter in the plain text can ever be represented in cipher by itself, e.g. N can never be represented by N.
8. Each single letter in the plain text can be enciphered by one of only five other letters; that immediately below it in the same column, and one of the four other letters in the same row.
9. Because of Rule 8, the 3 or 4 letters having the highest frequency in the cryptogram will usually be found in the same row of the matrix as E.
10. Two reversed digraphs in clear text, such as ER and RE, will always be enciphered by two reversed digraphs in the cipher text.
11. Of the 24 possible digraphs which it is possible to form by combining any one letter with the other 24 letters in the square matrix;

    16 are encipherable by using the letters in the opposite corners.

    4 are enciphered by letters in the same row,

    4 are enciphered by letters in the same column.

In the case of diagonal substitution only is encipherment reciprocal, e.g.

ER plain equals HA cipher, and HA plain equals ER cipher.

12. An equivalence such as DE equals EF necessarily implies the arrangement of D, E, and F, either DEF in the same line or D / E / F in the same column.

13. Two equivalences, such as BF equals ND and BM equals NI, indicate that B and N are in the same line.

14. Two equivalences, such as BY equals RU and BT equals NU, indicate that B and U are in the same column.

15. Two equivalences, such as UC equals PY and UH equals YD, indicate that U and Y are in the same line or in the same column, and the same for H and D.

16. It may be remarked that no change takes place in the equivalences of the digraphs of the plain text and those of the cipher text if we transpose the lines of the square matrix <u>in circular permutation</u> from top to bottom or the columns from left to right. There are 25 squares, beginning with the original one, that may serve for encipherment and yield identical cryptograms for the same plain text.

With the three diagrams below, we will try the word INFANTRY.

B A N K R	D E F G H	N K R B A
U P T C Y	I L M O Q	T C Y U P
D E F G H	S V W X Z	F G H D E
I L M O Q	B A N K R	M O Q I L
S V W X Z	U P T C Y	W X Z S V
Plain: IN FA NT RY	IN FA NT RY	IN FA NT RY
Cryptogram: MB EN TF YH	MB EN TF YH	MB EN TF YH

If we permute the lines or columns of the original square at random, the equivalents of the plain text digraphs enciphered diagonally are not changed, but the clear text digraphs enciphered in the same line or in the same column have different equivalents in cipher; thus:

B A N K R	N B A K R
U P T C Y	M I L O Q
D E F G H	F D E G H
I L M O Q	W S V X Z
S V W X Z	T U P C Y
Plain: IN FA NT RY	IN FA NT RY
Cryptogram: MB EN <u>TF</u> <u>YH</u>	MB EN <u>MN</u> <u>QR</u>

The Four Square Matrix Cipher

Another digraphic system employs a four square matrix, an example of which is illustrated in Figure 48. Quadrants I and III contain the two alphabets of the plain text in normal order, while Quadrants II and IV contain the two cipher alphabets in key-word mixed sequence. The keywords for the cipher alphabets in this case are WASHINGTON SQUARE and SWEET LAND OF LIBERTY.

	Quadrant I	Quadrant II	
Plain 1	A B C D E F G H I K L M N O P Q R S T U V W X Y Z	W A S H I N G T O Q U R E B C D F K L M P V X Y Z	Cipher 1
Cipher 2	S W E T L A N D O F I B R Y C D G H K M P Q V X Z	A B C D E F G H I K L M N O P Q R S T U V W X Y Z	Plain 2
	Quadrant IV	Quadrant III	

FIGURE 58

The letters of the plain text are written in groups of two, thus - the word ATTACK would be AT TA CK. The first letter of a pair is found in Plain 1 (Quadrant I) and the second in Plain 2 (Quadrant III); A is in Quadrant I and T in Quadrant III. These two letters are in opposite corners of a rectangle, their cipher equivalents are located in the opposite corners of this rectangle in Quadrants II and IV. (Cipher 1 and Cipher 2 squares), respectively, and are H and D. The word ATTACK will be enciphered as follows:

Plain - AT TA CK

Cipher - HD DT ID.

Decipherment is exactly the reverse of encipherment, the cryptogram being written in pair of letters and the cipher equivalents located in Quadrants II and IV, then the clear pairs will be found in the other corners of the rectangle in Quadrants I and III respectively.

Chapter XVI

CRYPTOGRAMS IN FOREIGN LANGUAGES;

CONCLUDING REMARKS

The methods of cryptanalysis described in the preceding chapters dealt entirely with messages the plain texts of which were in English. They are equally applicable to cryptograms when the clear texts are in other languages. The procedure is identical, but the frequency tables and other data on the language suspected will be used in place of the like material on English.

Although the ability to read the language of the original text is very helpful, it is quite possible for a person unfamiliar with French, German, Italian, or Spanish, to solve cryptograms the plain texts of which are in these tongues.

In analyzing cipher texts of messages transmitted by wire or radio channels, it is well to remember that the word STOP is frequently employed in telegrams in English instead of the period to separate sentences or other divisions of a message. The French use the same word, while the German equivalent is PUNKT and the Spanish PUNTO. The Italian is also PUNTO.

The principles involved in the solution of monoalphabetic substitution form a basis for attacking polyalphabetic ciphers, although additional means, not required for solving simple substitutions, will also be brought into play. Cryptanalysis of systems such as the Vigenère, Porta, Beaufort, and modifications of these, as well the digraphic ciphers of the Playfair and 4-square matrix types, is reserved for an advanced text and course.

Sir Francis Bacon's Rules for Ciphers

In his "OF THE ADVANCEMENT OF LEARNING," (Original Edition, 1605), Sir Francis Bacon referred to ciphers in the following terms:

> "But the vertues of them, whereby they are to be preferred, are three:
>
> that they be not laborious to write and reade; that they bee impossible

to discypher; and, in some cases, that they bee without suspition."

"But in regarde of the rawnesse and unskilfulnesse of the handes, through which they passe, the greatest matters are many times carryed in the weakest cyphars."

Fundamental Requirements for a Military Cipher

Kerckhoffs, in 1883, laid down the following requirements for a military cipher, (from "Manual for the Solution of Military Ciphers," by Colonel Parker Hitt):

1st. The system should be materially, if not mathematically, indecipherable.

2nd. It should cause no inconvenience if the apparatus and methods fall into the hands of the enemy.

3rd. The key should be such that it can be communicated and remembered without the necessity of written notes and should be changeable at the will of the correspondents.

4th. The system must be applicable to telegraphic correspondence.

5th. The apparatus should be easily portable and a single person should be able to operate it.

6th. Finally, in view of the circumstances under which it must be used, the system must be an easy one to operate and demand neither mental strain nor remembering a long series of rules.

It must be assumed that the enemy has knowledge of the general system and it is therefore necessary to rely for secrecy upon the key only.

Since the system must be applicable to telegraphic correspondence, this automatically eliminates all methods that do not use letters or Arabic figures. In order not to overload channels of communication, a cipher system must not prod secret text more than twice as long as the plain text.

General Givierge said that the purpose for which it is to be used and the circumstances under which it must be employed are the prime factors in the choice of a cipher system.

APPENDIX A

Glossary

ACROSTIC	A poem, or other composition, in which initial or other letters, taken in order, form a word, phrase or sentence.
ALPHABET	The arrangement of letters or characters by means of which a text is written.
" , Cipher	The alphabet of the cryptogram or secret text.
" , Deciphering	The cipher alphabet arranged for converting the cryptogram back into plain text.
" , Enciphering	The cipher alphabet arranged for converting plain text into secret text.
" , Mixed	An alphabet in which the elements have been re-arranged in a sequence not the usual order. The arrangement may be according to a definite system or at random.
" , Normal	The alphabet as customarily used to write plain texts in a language.
" , Direct Standard	The alphabet arranged in the direction of the order of letters as customarily written. It may begin with any letter, as long as the sequence is in the usual direction; from left to right, or right to left as the case may be.
" , Reversed Standard	The alphabet arranged in the opposite direction from that normally used for the sequence, commencing with any desired letter.
ANAGRAM	A word, phrase, or sentence, formed by re-arranging the letters of a different word, phrase, or sentence.
ANAGRAMMING	Re-arranging the letters of a word, phrase, or sentence, or other elements of a text.
BIGRAM	See DIGRAPH.
BI-LITERAL	Consisting of, or relating to, two letters.
BLOCK	A design, usually a square or other rectangle, containing smaller subdivisions or cells.
CELL	A subdivision of a design, normally a square or other rectangle.
CIPHER	A cryptographic system which usually deals with the individual letters or pairs of letters (sometimes words) of a message.
CHART	A sheet giving information. See "Graph" also "Table."

CIPHER, Substitution	A cryptographic system in which the elements of the plain text, letters, pairs of letters, and sometimes words, are replaced in the secret text by other letters, or figures, or symbols.
CIPHER, Transposition	One in which the elements of the plain text, letters, pairs of letters, or words, are simply re-arranged in a different sequence or order.
CODE	A system of signals for communication by flags, telegraph, etc. A cryptographic system which deals usually with elements of plain text messages that are of varying length - letters, pairs of letters, syllables, words, phrases, or even whole sentences.
CODE, Enciphered	When a message, already in code is enciphered to provide additional security, the resulting text is termed ENCIPHERED CODE.
CODE, One-Part	One in which the plain text meanings are usually listed in alphabetical order, with their code groups in alphabetical or numerical sequence alongside them. Used in both encoding and decoding.
CODE, Two-Part	One consisting of two parts: The encoding section having the meanings usually in alphabetical order and the code groups (letters, words, or figures) in random sequence as if mixed up and drawn from a hat; the decoding section having the code groups arranged in alphabetical or numerical order and the meanings, therefore, in random sequence.
CODE, Random	See CODE, Two-part.
COMPONENT	A constituent part.
COMPONENT, Cipher	That part of a cipher alphabet (enciphering or deciphering) in which the elements of the secret text are found.
COMPONENT, Plain	That part of a cipher alphabet in which the elements (usually letters) of the plain text are found.
"CRYPT"	A popular, abbreviated form of the word CRYPTOGRAM.
CRYPTO	An abbreviated form occasionally used in place of the words CIPHER and CRYPTOGRAM, as well as CRYPTOGRAPHIC. Sometimes used to designate an element of the secret text or the cipher text itself.
CRYPTANALYSIS	The steps and operations performed in the application of the principles of cryptanalytics.
CRYPTANALYST	A person versed in the art of cryptanalysis.

CRYPTANALYTICS	That branch of cryptology which deals with the principles, means, and methods employed in the solution or analysis of cryptograms without having knowledge or possession of the system or key.
CRYPTANALYZE	To solve a cryptogram by cryptanalysis.
CRYPTOGRAM	A message in code or cipher. The secret text of a communication in writing.
CRYPTOGRAPH, (Noun)	A mechanical device, usually hand-operated, for cryptographing and decryptographing messages.
CRYPTOGRAPH, (Verb)	To convert a plain text into secret text.
CRYPTOGRAPHER	A person skilled in cryptography.
CRYPTOGRAPHY	The science which deals with all the means and methods for converting an intelligible message into the unintelligible or secret text, and for reconverting the unintelligible or secret form into the intelligible message by a direct reversal of the steps used in the original process.
CRYPTOLOGIST	A person skilled in cryptology.
CRYPTOLOGY	That branch of knowledge which treats of all the methods of secret intercommunication. The means may be written, audible, visual, invisible, etc.
CRYPTONYM	A secret name.
DECIPHER	To reconvert the cipher text of a message into its plain text by a direct reversal of the steps used in enciphering the message.
DECIPHERER	The person who deciphers a cryptogram in cipher.
DECIPHERMENT	The result, also the process, of deciphering.
DECODE	To reconvert the code text of a message into its plain text. Also used as a noun to designate the clear text as decoded.
DECODER	The person who decodes a message in code. The term CODE CLERK is more generally used to indicate a person whose duty it is to encode or decode messages. CRYPTOGRAPHIC CLERK is preferable as it includes persons engaged with both codes and ciphers, or either.
DECRYPT	To solve a cryptogram by cryptanalysis.
DECRYPTMENT	The result, also the process, of cryptanalyzing a cryptogram.
DECRYPTOGRAPH	To reconvert a cryptogram into its plain text by a direct reversal of the steps used in cryptographing the message.

DIGR

DIGR

ENCI

ENCI

ENCI

ENCO

ENCO

FACT

FACT

FREQ

GRAP

HOMO

HOMO

IDIO

IDIO

INSO

INTR

KEY

DIGRAM — See DIGRAPH.

DIGRAPH — A pair, or group of two, letters, figures or symbols. E.g. - AL, 23, Ø1, or #? .

ENCIPHER — To convert the plain text of a message into secret form by means of a cipher system.

ENCIPHERER — A person who enciphers messages. A cryptographic clerk whose duty it is to encipher messages.

ENCIPHERMENT — The result, also the process, of converting a plain text into cipher text.

ENCODE — To convert a plain text message into secret form by means of a code system.

ENCODER — The person who encodes a message. A cryptographic clerk who encodes messages.

FACTOR — (Noun). Any of the elements, quantities, or symbols, which, multiplied together, form a product. More specifically - a quantity or number which is an element in determining the dimensions of a geometric figure in transposition, or the intervals between repetitions in substitution and the resultant cycle (period or key-length).

FACTORING — Determining the dimensions of a geometric figure in a transposition cipher or the key-length from the intervals between repeated groups in a substitution cipher.

FREQUENCY — The number of times that letters, figures, symbols, (elements of a text) occur, singly or in groups of two or more.

GRAPH — A chart or diagram showing information in graphic form of relationship between elements, or of the frequency with which elements occur. See Table, Frequency.

HOMOPHONES — See Variants.

HOMOPHONIC — Pertaining to, or partaking of the character of, homophones.

IDIOMORPHIC — Having a form of its own.

IDIOMORPH — A pattern-word or phrase.

INSCRIPTION — The process of writing in elements of or a complete text. The first step in certain processes of transposition.

INTERVAL — The distance between the repetitions of groups or sequences; the first of successive occurrences of a group being counted and the one next following disregarded.

KEY — That portion of the agreement between correspondents which controls the steps to be followed in cryptographing.

KEY, Literal	A key consisting of a word, phrase, sentence or longer sequence.
KEY, Numerical	One composed of a number or series of numbers, usually derived from a literal key.
LANGUAGE	See "Text."
MATRIX	That which gives form to the content enclosed within it. Specifically: A square, rectangle, or other geometric design in which cryptographic elements are inscribed, e.g. the square for the alphabet of the Playfair cipher.
MONOALPHABETIC	Involving or pertaining to a single alphabet.
MONOGRAM	A single written character.
MONOGRAPHIC	Relating to single letters, or characters.
MONOLITERAL	Involving or pertaining to single letters.
NULLS	Non-significant characters (letters, figures, or symbols) or words inserted in the text of a message to disguise it and render more difficult the task of the cryptanalyst or inserted as fillers to complete a group or pattern.
PALINDROME	A word, phrase or sentence which reads the same forward or backward. E.g. "DEED," and "NOON."
PATTERN	A design or form.
PATTERN-WORD	Words having a definite and distinctive pattern or design in the sequence or position of letters in the words themselves are called pattern-words, or "idiomorphs." E.g. ARTILLERY, ATTACK, DIVISION, AGAIN, RECEIVE, BEGINNING.
PENTAGRAM	A group or sequence of five characters; letters, figures, or symbols.
POLYALPHABETIC	Involving or pertaining to two or more alphabets.
POLYGRAM	A group or sequence of two or more characters.
POLYGRAPHIC	Involving or pertaining to groups of two or more characters.
POLYLITERAL	Involving or pertaining to two or more letters at a time.
SEQUENCE	The order in which anything occurs or appears. Also a series of two or more characters.
SUBSTITUTE	To put in the place of a person or thing.
SUBSTITUTION, Monoalphabetic	A substitution system in which a single alphabet is used to encipher the plain text.
Polyalphabetic	A substitution system employing two or more alphabets in enciphering the plain text. E.g. The Vigenere.

SYMBOLS — Characters that are neither letters nor figures; used in substitution ciphers.

SYSTEM, DIGRAPHIC — One in which encipherment is accomplished two characters at a time; e.g. The Playfair.

SYSTEM, General — That portion of the agreement between correspondents which constitutes the general or basic method of the cipher system, usually fixed and invariable. The steps followed in the process are controlled by the key.

SYSTEM, MONOGRAPHIC — One in which encipherment is one letter at a time.

SYSTEM, Polygraphic — One in which the plain text is enciphered two or more letters at a time.

SYSTEMS, Monoalphabetic, Polyalphabetic, etc. — See "Substitution, Monoalphabetic," etc.

TABLE, Single Frequency — One showing the frequency of occurrence of single (individual) characters.

TABLE, Digraphic — A chart showing the frequency of occurfrence of digraphs.

TABLE, Trigraphic — One showing the frequency of trigraphs.

TETRAGRAM — A group or sequence of four characters.

TEXT, Clear, or Plain — The text of a message that, on its face, conveys an intelligible meaning in the language in which it is written. Also called "Plain Language" and "Clear."

TEXT, Cipher, or Secret — The text of a message, that, on its face, conveys no intelligible meaning in the language in which it is written. The Cryptogram.

TRANSCRIPTION — Writing a copy of. Specifically: the second step in certain processes of transposition.

TRANSPOSE — To change the position, order, or sequence of anything.

TRANSPOSITION — Changing the position, order or sequence of anything. One of the two main classes of ciphers.

TRIGRAM — See "Trigraph."

TRIGRAPH — A group or sequence of three characters.

TRIGRAPHIC — Relating to trigraphs, (groups of three characters); as monographic relates to one, digraphic to two.

TRI-LITERAL — Involving or pertaining to three letters.

VARIANTS — Two or more characters representing the same plain text letters. Homophones.

APPENDIX B

Bibliography

The following is a brief list of works published in English. Books in foreign languages are not generally available. For a comprehensive and annotated bibliography consult "Traité de Cryptographie," by Andre Lange and E.-A. Soudart; Librairie Felix Alcan, Paris; 1925 and 1935.

d'Agapeyeff, Alexander — "Codes and Ciphers," London; Oxford University Press; 1939. pp. 158-2. (Meridian Books).

Candela, Rosario — "The Military Cipher of Commandant Bazeries," New York; The Cardanus Press; 1938. pp. xv., 137.

Gaines, Helen Fouche — "Elementary Cryptanalysis," Boston; American Photographic Publishing Co.; 1939. pp. vi., 230.

*Hitt, Parker — "Manual for the Solution of Military Ciphers," Leavenworth, Kansas; Press of the Army Service Schools; 1916 and 1918. pp. 101.

Lysing, Henry — "Secret Writing," New York; David Kemp and Company; 1936. pp. x. - 118.

*Langie, Andre — "Cryptography," Translated from the French by J. C. H. Macbeth, London; Constable & Co., Ltd., 1922. pp. vii. - 192.

Mansfield, Louis C. S. — "The Solution of Codes and Ciphers," London; Alexander Maclehose & Co.; 1936. pp. xii. - 162.

Ohaver, M. E. — "Cryptogram Solving," Columbus, Ohio; Stoneman Press; 1933. pp. 32. (Pamphlet).

Pratt, Fletcher — "Secret and Urgent, the Story of Codes and Ciphers," Indianapolis; The Bobbs-Merrill Company; 1939. pp. 282. Reprinted 1942, by Blue Ribbon Books, New York.

Thomas, Paul Bernard — "Secret Messages, how to read and write them," New York; A. A. Knopf; 1928. pp. 4 - 138.

White, Maurice — "Secret Writing," Washington, D.C.; The Washington Service Bureau; 1938. pp. 24. (Service Booklet No. 181.

*Yardley, Herbert Osborn — "The American Black Chamber," Indianapolis; The Bobbs-Merrill Company, 1931, pp. 375. Reprinted by Blue Ribbon Books, New York.

*Note. These works are out-of-print and rare. Those by Mansfield and Thomas are rather difficult to find. The others should be available.

APPENDIX C

Table 1

DATA ON THE ENGLISH LANGUAGE
(Compiled from various sources)

A. FREQUENCY OF SINGLE LETTERS

1. Literary Text - Count of 200, reduced from total of 272,000.

A	B	C	D	E	F	G	H	I	J	K	L	M
16	3	6	8	26	5	4	11	14	Ø	1	8	5

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	16	4	Ø	12	13	18	6	2	4	Ø	4	Ø

2. Telegraph Text - Count of 200, reduced from 10,000 (Hitt).

A	B	C	D	E	F	G	H	I	J	K	L	M
16	3	6	8	26	4	4	8	14	1	2	8	6

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	17	5	1	14	13	13	6	3	3	1	4	Ø

B. ORDER OF FREQUENCY, SINGLE LETTERS

1. Literary Text - E T A O I N S R H D L U C F M P Y G W B V K J Q X Z.
2. Telegraph Text - E O A N I R S T D L H U C M P Y F G W B V K X J Q Z.

C. ORDER OF FREQUENCY, Miscellaneous

1. Digraphs - TH ER ON AN RE HE IN ED ND HA AT EN ES OF OR NT EA TI TO IT ST IO LE IS OU AR AS DE RT VE.
2. Trigraphs - THE AND THA ENT ION TIO FOR NDE HAS NCE EDT TIS OFT STH MEN.
3. Double Letters - SS EE TT LL MM OO FF.
4. Initial Letters- T O A W B C D S F M R H I Y E G L N P U J K.
5. Final Letters - E T D S N R Y F L O G H A K M P U W.
6. 1-Letter Words - A, I.
7. 2-Letter Words - OF, TO, IN, IT, IS, BE, HE, BY, OR, AS, AT, AN, SO, WE.
8. 3-Letter Words - THE, AND, FOR, ARE, BUT, ALL, NOT, WAS, YOU, GET, SEE.
9. 4-Letter Words - THAT, WITH, FROM, HAVE, THIS, THEY, WILL, WANT, KNOW.

D. SPECIAL CHARACTERISTICS

1. Q is always followed by U.
2. J and V are always followed by a vowel.
3. The percentages are:
   (a) Literary Text - E, 12.77%; vowels A E I O U, 38.37%; consonants L N R S T, 31.86%; consonants J K Q X Z, 1.77%; and the remaining letters - 28%.
   (b) Telegraph Text - E, 13.19%; vowels A E I O U, 40.08%; consonants L N R S T, 30.77%; consonants J K Q X Z, 2.25%, and the remaining letters - 26.9%,
   (c) Approximate figures are: E, 13%; A E I O U, 40%, L N R S T. 30%; J K Q X Z, 2%, remaining letters, 28%.

Table 2

First 600 letters taken from "Kontentment" - one of "Josh Billing's Works

A
DKVDHWSEHSMLHLGEGMWMHHHHNMED LMZRLHWMEFHNNNNMWERIH
SNTLRCMZ PC LMPYRVTNSOPOVNMNNC MRKLZRNNNS NNNMNNRCDA

B
NEOTWDGTW
EUEEC IUEE

C
EAAATSUKAKNN
UQKLOOKEKOOO

D
OERENIAELALELENNLLEA
AATTLOENFBNEWHAOMDOM

E
TMHBTLTRTMSBNMNHNSTNB LVDTBTTHHKRDHHHPFTTHRTMNRTTMBRC LHHTRKTNRETMNHTBN
NNGCDBWKNNOTSASSS WNTLATTWKNDHWRONORNRKNDAFNNSWNNNAOSWNDNHASSENNNOIWHS

F
ISDIUFFREIOS
FOOHFFOEUTAU

G
EVNNNNNNV
IOAIBLIMR

H
TSTZTTVTTRTFENTVTTWDTTTTWOTETHTE
EAEAIEAOAAIEAEAAEEEEAI IAEEIEHIEA

I
TSTLTTUTHYWRYHRBGZTKZSTHMGHELWHSTTETSMTP
STVTTSRSSNTDNSFNTNZNNFNSZTNZKTSNNZNNATTN

J

K
-TECENEYATCICN
QAOOOETOIOCECT

L
UATALDERURTCALUGUL
TIELAOEDDOYALADLLO

M
TTASATSAATVTDYRAGTD
EEEAOAAAAEOEAOIAOEI

N
OFEAOEEIUUI IUENWOAOEDI IDAEEAOEAIOEEI IAENI OEE IADAAAAAEAOAENI OII EROEEOI II
TTTBTTTETTEGDNYETWTTOTHTKUAGTTDSTTTDETNYGTTTTWTYAAGAGTTNYGWTECCTTTESKE

O
KTGKTKFMTRLOHYDKNWFWFEWKCCDMKWWMWSDMCOLCCZE
NVDNBURUVUOSUWNNRRRURTRNUNVRNRURRHNRONVRNNV

P
APAPAPSAP
PI FI PYEPI

Q
C
U

R
AIOAEOOOEOEOEAOAOOOOAGO
OETOIHLIETWLFEAELEMEAEN

S
AII IEESITEESOAITUNESIAESTSIEGNTES
ITHSASTAFMALEMWPTCSIWUSTSOITFTIS-

T
NNNZFLATUIDNNNDESNRNIUNEAENNNNNROIKNNSNNVFNYUNNNNNSINNYINSNNNVSKIT
EMIHOIELIIEEMIOHHSHOHEBTMEEEHHHHSLEECEMIHEHKEMIHBHSWEHHEIEMITISTE

U
CBQOOOOONFOBSOF
LTINNNTLFSZTCLL

V
OIOAAOOO
GAHEHMGT

W
EEONESERDESNEZNE
AINABOWOHAOOOHIB

X

Y
ANPLNNN
IOIKTMT

Z
AUIAIII
HIAIWLO

Prepared by Alvin Cooperman, Tuesday-Thursday Class - Cryptography 1, Washington Square College, New York University, October, 1942; Donald D. Millikin, Major, Retired, Instructor.

Table 3

First 600 letters taken from a short story "Smart Dope" by Jerome Weidman

A
-TWSHELHWSLMOCTESELESHFMTSDSDHERTLSXNNTWHTOMEEWHOLSCMC
NLSYPRWTSLBYLNNRLDSRYLSRSRLYHDFLJWNMTPWYTBSRNSSVTWLYNK

B
AYTAOS
OEUOEE

C
SUSSSUE
HAHHHAA

D
NEEINNEAIDNLAILLU
YHINEITYDLAAMENNO

E
THYMNPNYHYBDHHYRHHLYYVHHOHMHHRSSTKNPNVHDHSSHRHIHHNNOTVVBTLNBRKH
NDYYENPSFYDTGFSSDMOSBNDHDYDIDDNSTICTBNOHKNTTIDKDDSSDBNYILPIFYTC

F
ISEOLAE
IIIHATO

G
INIENIUITU
HNHOHHHHOO

H
OOGTOTDCLWGTSYTFNYSTGSSACTEGGTWSGTTCWSKMTNCPT
IETIIAIOEATEEEEIEAEEEEDAOEERTAYOTEEOYOIEEAOOE

I
LHRTFHHYHFROTRDFNSMHTYELELMSTSRTRRIWMTWHTNEWO
SMGFTNMTSRNDTGTNSNDSMNSTTTNTODGTOGNTLETMOTNTN

J
A
O

K
LTOELESOA
SNENNNHNE

L
UANOALALADSNAALUALSOUOUINOEALU
IKAHLALRREAYFLTOLIAIDKDTATELY-

M
IOITESIDATISN
TESAIAEDIIHAA

N
AEIGEIDKAUAIIEOUIENAIOOEOKEDIKDKOAOOEIKIOA
DTGEELTOUDDITHLDGSOEAAEDTOOTTOTOEEITHLOTEM

O
YTSSTHONBSYWESGRHMNJHOISNENNISHBHONHNGTITOYDTHOTNYGTHFYY
UHHMHOLWUAUYFNEUURTKOLNNNUWUUWUYOLWUWANNOYUNBOLKWUIANRUU

P
APAE
PEEH

Q
Y
U

R
SIALEALAAAOEEHEYEAO
ISIISSESTOEASIIIETE

S
IKEASMIRWAERIREIAREASSEEIEENTTWTRTWTSUSYSYAEEAEAWTY
TORSOAFSCAOTHAILTAOSMHAAHHEENLCAOIESESOHCHKTMYTTCAB

T
SNMHEIEWISNAIUEHSIUDNSRLEFIOIIANUHYAINHTNINLISRSISNELINAIE
EOAYIWHOHYKIWYMYABHWHYAHIESASSIHSAHISATHHWMIGHOIWOEOOYOESH

U
OOONBOOOOOOOQOOOO
LTCNTNLTGSLLEDGCL

V
EEEA
EEEE

W
TEAOWTTAODOOTOTAO
ATSWHAOSSAHIHIASI

X
E
A

Y
DTATTATODTALAAHOHSOTAEL
OEIEEBOEHEEHITSRSQOOSOO

Z

Prepared by Alvin Cooperman, Tuesday-Thursday Class, Cryptography 1, Washington Square College, New York University, October, 1942; Donald D. Millikin, Major, Retired, Instructor.

Table 4

These trigraphs taken from first 602 letters of Shakespeare's "Hamlet"

A	NYTDNCRHDIKEHHUTROMUWMHETDDFHNHLLNSHRARMHHGEDGHRSFT YNNNRRNNNMTRVDRMTNRLTKSRNNNRTRSCBRYTTPTRTSPRGIVTYNS	51 7
B	GOSHUAS EEIIEEU	15
C	UTCNSURIRTAELLR OAKIOHOKEHEEOOE	25
D	NNLRELNAROUNINNNNEOLVRROR AUYOFAIQNNOMTHSALANIYOOMA	64
E	HRWMSVHBHMRWVGHEBRITHVISWMECHNKTHHHRIGMHNVRWNIRIBCBHRICWMWMCPVSE RNRSLTRRYMFLGTETDLFRAYTSLETLRMHIAMRFGMNDGYWLSRLVRHRRACOLALGLASSN	12
F	NLEDOEIOEOOR OLUROMYMRAHA	16
G	NNETNLIEEUIEINAN LBEVWOHEIOHOHAAH	46
H	TTORTTCTTTUGTTCTETITDWTITTGLWOTOEWSTFEGWTTTITG EEEOEITAEAATOEBEAIEEOOENIETOOARAOAOEIOTAAIIAIO	41
I	LKTCHLTBDSUTRNTRBEHKORHLGNDLTTPHTNHHAOHTT VNSSSEDTACERNGOVDTNHSESEVGEESOEMOGSNNINOS	Ø
J		6
K	ECNCAN ITSAEI	25
L	OEFGULEEOELELADELOPOLEEEL DFOILYVIDLGLUSILHDALACCLU	18
M	ROEFAAODFEMEESIOOD EEOUSOEAYMASEYWEEA	35
N	EAAAUBIROEAAADIDAIAEUAEADORRNIIENIA ASDDFGGAYOCKDOGIDKDDDDTEIEAAIGTNOGT	56
O	HFYLDYCMPYHNTCFEYNMGOYDHISHHTRTYGOTHSHYDHDHIECHICGOTNHIT SLUNHUMSNUUWBFRLUTUODUMRAFWITUTUODFNLHUHLSRTFMRWMODITRSU	5
P	UYAAP OLIPE	1
Q	D U	38
R	EEUEAAUUTFOSEAAIROAEAEFGAEHEAEAOEOAAOU EMSNDEHTUATECTDRIACITEIOEWENONDAEACOAF	34
S	ONEROIWIIKIMUEULAMIDIETSOIUAIEOYIA TWTETNTCRTBIETTOTTTTGTOMAHNTTEATBY	52
T	SSESRSKETERHSITAREOSHEASADSIRSSSONOHSAAAOAHASSNOASUN HAHCIRWTHOHHITEHHGAIIHIHCHEHHAHOHOHOSHIIHIWHHHOHIIQA	21
U	DOOFYOORMOQGOOLOOOLBO NRCLPRRCCHIASDSNGBSTR	7
V	ILAIIEA EEEAEDE	13
W	-SOTGYOERYMOT OESEEAHEHHEEA	Ø
X		14
Y	ADELNEFMEDMAAS AOOUOOONOOPWS-	Ø
Z		

Prepared by Alvin Cooperman, Tuesday-Thursday Class - Cryptography 1, Washington Square College, New York University, October, 1942; Donald D. Millikin, Major, Retired, Instructor.

Table 5

"KONTENTMENT"			"SMART DOPE"			"HAMLET"		
N	71	11.8%	E	63	10.5%	E	664	10.7%
E	70	11.7%	T	58	9.7%	O	56	9.3%
T	65	10.8%	O	56	9.3%	T	52	8.7%
A	48	8.0%	A	54	9.0%	A	51	8.5%
O	43	7.2%	S	51	8.5%	H	46	7.5%
I	40	6.7%	I	45	7.5%	I	41	6.8%
S	33	5.5%	H	45	7.5%	R	38	6.4%
H	32	5.3%	N	42	7.0%	N	35	5.8%
R	23	3.8%	L	30	5.0%	S	34	5.7%
D	20	3.3%	Y	23	3.8%	L	25	4.2%
M	19	3.2%	R	19	3.2%	D	25	4.2%
L	18	3.0%	D	17	2.8%	U	21	3.5%
W	16	2.7%	U	17	2.8%	M	18	3.0%
U	15	2.5%	W	17	2.8%	G	16	2.7%
K	14	2.3%	M	13	2.2%	C	15	2.5%
C	12	2.0%	G	10	1.7%	Y	14	2.3%
F	12	2.0%	K	9	1.5%	W	13	2.2%
G	9	1.5%	F	7	1.2%	F	12	2.0%
P	9	1.5%	C	7	1.2%	V	7	1.2%
B	9	1.5%	B	6	1.0%	B	7	1.2%
V	8	1.3%	P	4	.7%	K	6	1.0%
Y	7	1.2%	V	4	.7%	P	5	.8%
Z	7	1.2%	J	1	.2%	Q	1	.2%
Q	1	.2%	Q	1	.2%	J	Ø	
J	Ø		X	1	.2%	X	Ø	
X	Ø		Z	Ø		Z	Ø	
	600			600			600	

Prepared by Alvin Cooperman, Tuesday-Thursday Class - Cryptography 1, Washington Square College, New York University, October, 1942; Donald D. Millikin, Major, Retired, Instructor.

Table 6

DATA ON THE FRENCH LANGUAGE
(Compiled from various sources)

A. FREQUENCY OF SINGLE LETTERS - Count of 200, reduced from a total of approximately 25,000.

A	B	C	D	E	F	G	H	I	J	K	L	M
15	2	7	9	35	2	1	1	14	Ø	Ø	12	6

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
17	13	6	1	14	14	13	14	3	Ø	1	Ø	Ø

B. ORDER OF FREQUENCY

1. Single Letters - E N A S R I U T O L D C M P V F B G X H Q Y Z J K W.
2. Digraphs - ES EN LE DE ON OU NT RE NE ED TE EM SE ER AR ME AN IT ET IE TI EL NS UR.
3. Trigraphs - ENT EDE LES LLE QUE AIT EME ION EUR ELL SSE EST DAN DEL MEN DES TIO ESE ANS TER ONS QUI AIS OUS.
4. Double Letters - SS LL EE NN TT FF CC RR MM PP.
5. Initial Letters- D L E P A C S M R I F.
6. Final Letters - E S T R A N L I U D C X.
7. 1-Letter Words - A, Y, O.
8. 2-Letter Words - AN, AU, CE, CI, DE, DU, EN, ET, IL, JE, LA, LE, MA, ME, NE, NI, ON, OU, SA, SE, SI, UN.
9. 3-Letter Words - LES, QUE, PAR, EST.

C. SPECIAL CHARACTERISTICS

1. Q is always followed by U in the body of words.
2. X is generally preceded by U, sometimes by another vowel.
3. K and W occur rarely and usually in proper names or foreign words.
4. Z is preceded by E; H occurs in the trigraph CHE and is generally preceded by C, P, or T.
5. The percentage of E is 17.5; vowels A E I O U comprise 45.5%, consonants L N R S T - 35%, consonants J H X Y Z - 2%, and the remaining letters approximately 17.5% of the total.

Table 7

DATA ON THE GERMAN LANGUAGE
(Compiled from various sources)

A. FREQUENCY OF SINGLE LETTERS - Count of 200, reduced from a total of approximately 36,000.

A	B	C	D	E	F	G	H	I	J	K	L	M
10	4	5	10	36	3	7	9	15	1	3	6	5

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
20	6	1	Ø	15	13	12	10	2	3	Ø	Ø	4

B. ORDER OF FREQUENCY

SINGLE LETTERS - E N R I S T U D A H G L O C M B Z F W K V P J Q X Y

DIGRAPHS - EN ER CH DE GE EI IE IN NE BE EL TE UN ST DI ND UE SE AU RE HE IT RI TZ.

TRIGRAPHS - EIN ICH DEN DER TEN CHT SCH CHE DIE UNG GEN UND NEN DES BEN RCH.

TETRAGRAPHS - ICHT KEIT HEIT CHON CHEN CHER URCH EICH DERN AUCH SCHA SCHE SCHI SCHO SCHU.

DOUBLE LETTERS - EE TT LL SS DD MM NN.

FINAL LETTERS - N E R T S D H U Z F, etc.

FINAL DIGRAPHS - EN ER, etc.

2-LETTER WORDS - AB AM AN DA ER ES OB SO WO IM IN UM ZU DU JA.

3-LETTER WORDS - DIE DER UND DEN DAS ICH VON IST DES SIE DEM MIT EIN BEI MAN NUR WAR IHN ZUM ZUR.

4-LETTER WORDS - DASS SICH EINE ABER WENN, etc.

C. SPECIAL CHARACTERISTICS

1. Q, X and Y occur rarely.

2. Q is always followed by U. C is usually followed by H or K.

3. The percentage of E is 18; vowels A E I O U comprise 37.84%, consonants L N R S T - 33.01%, consonants J K Q X Z - 3.72%, and the remaining letters 25.43% of the total.

Table 8

DATA ON THE ITALIAN LANGUAGE
(Compiled from various sources)

A. FREQUENCY OF SINGLE LETTERS - Count of 200, reduced from a total of approximately 23,000.

A	B	C	D	E	F	G	H	I	J	K	L	M
21	2	9	7	26	2	4	2	23	Ø	Ø	13	5

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	18	6	1	13	12	12	6	3	Ø	Ø	Ø	2

B. ORDER OF FREQUENCY

1. Single Letters - E I A O R L N T S C D P U M G V H Z B F Q J K W X Y.

2. Digraphs - ER ES ON RE DE DI TI SI EL EN LA AL NT RA CO TA TO LE LI AN IN IO AR OR.

3. Trigraphs - CHE ERE ZIO DEL QUE ARI ATO ECO EDI IDE ESI IDI ERO PAR NTE STA MEN.

4. Double Letters - LL SS TT EE PP NN BB GG CC.

5. Final Letters - I A E O N L R D U.

6. 1-Letter Words - E A I O.

7. 2-Letter Words - DI IN HA HO.

8. 3-Letter Words - CHE PER CON DEL QUE.

C. SPECIAL CHARACTERISTICS

1. J, K, W, X and Y occur rarely and, as a rule, only in proper names.

2. Q is always followed by U.

3. H is most frequently preceded by C or G.

4. The percentage of E is 13; vowels A E I O U constitute 47%, consonants L N R S T - 31.5%, consonants J Q X Z - 1.5%, with the remaining letters forming 20% of the total.

Table 9

DATA ON THE SPANISH LANGUAGE
(Compiled from various sources)

A. FREQUENCY OF SINGLE LETTERS - Count of 200, reduced from a total of approximately 35,000.

A	B	C	D	E	F	G	H	I	J	K	L	M
26	2	9	10	28	2	2	2	13	1	ø	11	5

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
12	17	6	2	14	15	9	8	2	ø	ø	3	1

B. ORDER OF FREQUENCY

1. Single Letters - E A O S R I N L D C T U P M Y Q G B H F V J Z K W X.
2. Digraphs - ES EN EL DE LA OS AR UE RA RE ER AS ON ST AD AI OR TA CO SE AC EC CI IA.
3. Trigraphs - QUE EST ARA ADO AQU DEL CIO NTE OSA EDE PER IST NEI RES SDE.
4. Double Letters - EE LL RR AA SS CC DD NN.
5. Final Letters - A E S O N L R Y I D Z U.
6. 1-Letter Words - A E O U Y.
7. 2-Letter Words - EN LA DE LO EL SE.
8. 3-Letter Words - QUE DEL LOS.

C. SPECIAL CHARACTERISTICS

1. K, W and X occur rarely, usually only in proper names.
2. Q is always followed by U and the U by one of the other vowels, A, E, I or O.
3. H, J, V and Z are followed by vowels.
4. The percentage of E is 14; vowels A E I O U comprise 46%, consonants L N R S T - 30.5%, consonants J Q X Z - 2%, and the remaining letters 21.5% of the total.
5. A considerable number of words commence with LL.

Table 10

SOME LANGUAGE PECULIARITIES

1. French may be distinguished from English by the high frequency of Q and the low frequency of H, with K and W practically non-existent.
2. German may be distinguished from English by the high frequency of I, K, N, and W, and by the low frequency of A and O.
3. German may be distinguished from French by the high frequency of K and W and by the low frequency of A, L, and O.
4. German may be distinguished from Spanish by the high frequency of K and W, the unusually high frequency of N, and the low frequency of A and P.
5. Spanish may be distinguished from English by the high frequency of A and Q, the low frequency of T, and the absence of K and W.
6. Spanish may be distinguished from French by the high frequency of A and O and the low frequency of T and U.
7. Italian may be distinguished from English by the high frequency of A and I, the low frequency of H, and the rarity of J, K, W, X, and Y.
8. Italian may be distinguished from French by the high frequency of A and I, the low frequency of U, and the fact the frequency of E is lower and O higher than in French.
9. Italian may be distinguished from German by the high frequency of A, I, O, C, and L, and the low frequency or rarity of H, J, K, N, and W, together with the unusually higher frequency of P.
10. Italian may be distinguished from Spanish by the high frequency of I, the low frequency of Y and the absence of J.

Table 11

LANGUAGE DATA

MOTS "VIDES"

(Words normally omitted in telegraphic style).

ENGLISH	FRENCH	GERMAN	ITALIAN	SPANISH
A	À	An	A	De
An	De	Dass	Che	Del
And	Du	Den	Con	El
At	Est	Der	Di	En
For	Et	Die	E	La
Has	La	Es	È	Las
In	Le	In	Ha	Lo
Is	Les	Ist	Ho	Los
It	Par	Und	In	Que
Of	Que	Zu	Per	Se
On	Se	Ein		Y
Or				
The				
To				

APPENDIX D

Table 1

10-Letter Words for Constructing the Playfair Square

ADMIXTURES
ALPENSTOCK
AMPLITUDES
ANCHORITES
ANGLEWORMS
ARTICHOKES
AUTHORIZED
BACKGROUND
BACKSLIDER
BANKRUPTCY
BICHROMATE
BIFURCATED
BIGMOUTHED
BIVOUACKED
BOUNDARIES
BUCKINGHAM
CHIVALROUS
COMPATIBLE
COMPLAINTS
CUMBERLAND
DESOLATING
DESTROYING
DISCOURAGE
EMULATIONS
EXHAUSTION
EXHUMATION
FLOURISHED
FORMIDABLE
GELATINOUS
HYDRAULICS
HYPNOTIZED
HYSTERICAL
IMPERSONAL
IMPORTANCE
INOCULATES
ISOTHERMAL
JOURNALIST
JOURNALIZE
JUXTAPOSED
KALIGENOUS
LABYRINTHS
LACHRYMOSE
LANIFEROUS
LOGARITHMS
LONGCHAMPS
MACKINTOSH
MENDACIOUS
METAPHYSIC
MINERALOGY
MISFORTUNE
MODERATING
MONARCHIES
NEUROPATHS
OUTLAWRIES
OUTWEARING
PATRONYMIC
PECULATION
PLAYWRIGHT
PRECAUTION
PRESUMABLY
PREVIOUSLY
PROCLAIMED
PROFLIGATE
PROMULGATE
PURCHASING
QUATREFOIL
REFRACTION
REGULATION
REPUBLICAN
SPHEROIDAL
SUBJECTION
TAMBOURINE
TOURMALINE
ULCERATION
UNISOLATED
UNPOETICAL
UNSOCIABLE
VULCANIZED
WORKINGDAY
XYLOGRAPHS

Table 2

Prefixes

a-
ab-
ac-
ad-
af-
ag-
al-
an-
ana-
ant-
ante-
anti-
ap-
ar-
arch-
as-
at-
auto-
be-
bi-
circum-
co-
centi-
col-
com-
con-
contra-
cor-
counter-
de-
deca-
deci-
demi-
di-
dia-
dis-
em-
en-
enter-
es-
ex-
extra-
for-
fore-
hecto-
hemi-
hepta-
hetero-
Hexa-
holo-
homo-
hyper-
hypo-
ig-
il-
im-
in-
inter-
intra-
intro-
ir-
juxta-
mal-
male-
mega-
meta-
micro-
milli-
mis-
mono-
multi-
non-
ob-
oc-
off-
omni-
out-
over-
pan-
para-
pari-
pen-
per-
peri-
poly-
post-
pre-
pro-
proto-
pseudo-
re-
retro-
semi-
sin-
sine-
step-
sub-
subter-
suc-
super-
sur-
sym-
syn-
trans-
tri-
ultra-
un-
under-
uni-
up-
vice-
with-

Table 3

Suffixes

-able
-ac
-aceous
-acious
-acy
-age
-al
-ably
-an
-ana
-ance
-aneous
-ant
-ar
-arch
-ard
-art
-ary
-asm
-ast
-ate
-ble
-cal-
-cle
-cule
-cy
-d
-dom
-ean
-ed
-ee
-eer
-el
-elle
-en
-ence
-ency
-ent
-eous
-er
-erel
-erly
-ern
-ery
-es
-escence
-escent
-ese
-esque
-ess
-est
-et
-ette
-ful
-fold
-fy
-geneous
-fenous
-gram
-graph
-graphy
-head
-hood
-ian
-ible
-ibly
-ic
-ical
-ice
-ics
-id
-ide
-ie
-ier
-ies
-ile
-ine
-ing
-ion
-ious
-ique
-ise
-ish
-ism
-ist
-ite
-ity
-ive
-ix
-ize
-kin
-le
-lence
-lency
-lent
-less
-let
-ling
-lite
-logy
-long
-ly
-ment
-meter
-mony
-most
-nce
-ncy
-ness
-oid
-oidal
-on
-or
-ory
-ose
-ous
-phorus
-ple
-red
-ric
-ry
-s
-scope
-scopy
-ship
-sion
-some
-ster
-sy
-th
-teen
-ter
-ther
-ths
-tion
-tor
-tory
-trix
-tude
-ty
-ule
-ure
-ward (s)
-way
-ways
-wise
-y
-yte

Table 4

TWENTY-FIVE COMMONEST WORDS OF ENGLISH

Written Matter Dewey's Analysis	Words	Telephone Conversation Number of Conversations	Total Occurrence
the	I	467	3999
of	you	499	3540
and	the	496	3110
to	a	487	2060
a	on	458	2046
in	to	472	1942
that	that	397	1792
it	it	417	1605
is	is	419	1506
I	and	391	1363
for	*get	393	1360
*be	*will	402	1305
*was	of	396	1190
*as	in	408	1170
you	he	297	1115
*with	we	294	1100
he	*they	253	913
on	*see	328	887
have	have	367	883
by	for	330	823
*not	*know	338	753
*at	*don't	301	640
*this	*do	302	638
are	are	293	618
we	*want	297	599
		Total	36945

*Words not common to both lists.

The above table is taken from the article "The Frequency of Occurrence of Speech Sounds in Spoken English," by Norman R. French and Walter Koenig, Jr., published in the Journal of the Acoustical Society of America, 1929, page 113.

APPENDIX E

Table 1

FREQUENCY GRAPH

ENGLISH, TELEGRAPHIC TEXT - 200 LETTERS (HITT)

A - 16 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

B - 3 1 1 1

C - 6 1 1 1 1 1 1

D - 8 1 1 1 1 1 1 1 1

E - 26 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

F - 4 1 1 1 1

G - 4 1 1 1 1

H - 8 1 1 1 1 1 1 1 1

I - 14 1 1 1 1 1 1 1 1 1 1 1 1 1 1

J - 1 1

K - 2 1 1

L - 8 1 1 1 1 1 1 1 1

M - 6 1 1 1 1 1 1

N - 14 1 1 1 1 1 1 1 1 1 1 1 1 1 1

O - 17 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

P - 5 1 1 1 1 1

Q - 1 1

R - 14 1 1 1 1 1 1 1 1 1 1 1 1 1 1

S - 13 1 1 1 1 1 1 1 1 1 1 1 1 1

T - 13 1 1 1 1 1 1 1 1 1 1 1 1 1

U - 6 1 1 1 1 1 1

V - 3 1 1 1

W - 3 1 1 1

X - 1 1

Y - 4 1 1 1 1

Z - Ø

Table 2

INTERNATIONAL MORSE CODE AND CONVENTIONAL SIGNALS

To Be Used for All General Public Service Radio Communication

A .—	G — — .	N — .	U .. —
B — ...	H	O — — —	V ... —
C — . — .	I ..	P . — — .	W . — —
D — ..	J . — — —	Q — — . —	X — .. —
E .	K — . —	R . — .	Y — . — —
F .. — .	L . — ..	S ...	Z — — ..
	M — —	T —	

1 . — — — —	6 — ....
2 .. — — —	7 — — ...
3 ... — —	8 — — — ..
4 .... —	9 — — — — .
5	0 — — — — —

Ä (German) . — . —	É (French) .. — ..
Á or Å (Spanish-Scandinavian) . — — . —	Ñ (Spanish) — — . — —
CH (German-Spanish) — — — —	Ö (German) — — — .
	Ü (German) .. — —

Period	(.) . — . — . —	Distress Signal	... — — — ...
Semicolon	(;) — . — . — .	Attention Call (Obsolete Except in Amateur Work)	— . — . —
Comma	(,) — — .. — —	General Inquiry Call	.. — . — .
Colon	(:) — — — ...	From (de)	— .. .
Interrogation (Question Mark)	(?) .. — — ..	Invitation to Transmit (Go Ahead)	— . —
Apostrophe	(') . — — — — .	Wait	. — ...
Hyphen or Dash	(-) — .... —	Understand	... — .
*Signal Separating Whole Number From Fraction	. — .. —	Error	.....
		Received (OK)	. — .
Bar indicating fraction	— .. — .	End of Each Message (+)	. — . — .
Parenthesis	( ) — . — — . —	Transmission Finished (End of Work) (Conclusion of Correspondence)	... — . —
Quotation Marks	(" ") . — .. — .	& (Not Official) (ES)	. ...
Underline	.. — — . —	$ (" ") (SX)	... — .. —
Double Dash	(=) — ... —		
Break Signal			

1. A dash is equal to three dots.
2. The space between parts of the same letter is equal to one dot.
3. The space between two letters is equal to three dots.
4. The space between two words is equal to five dots.

*A mixed number, such as 2 2/25 is transmitted as 2 dit dah dit dit dah 2/25. The use of the separating signal prevents mistaking the mixed number for the fraction 22/25.

(Courtesy of R.C.A. Institutes, Inc., New York)

Table 3

CHART FOR PRINTING LETTERS AND NUMERALS

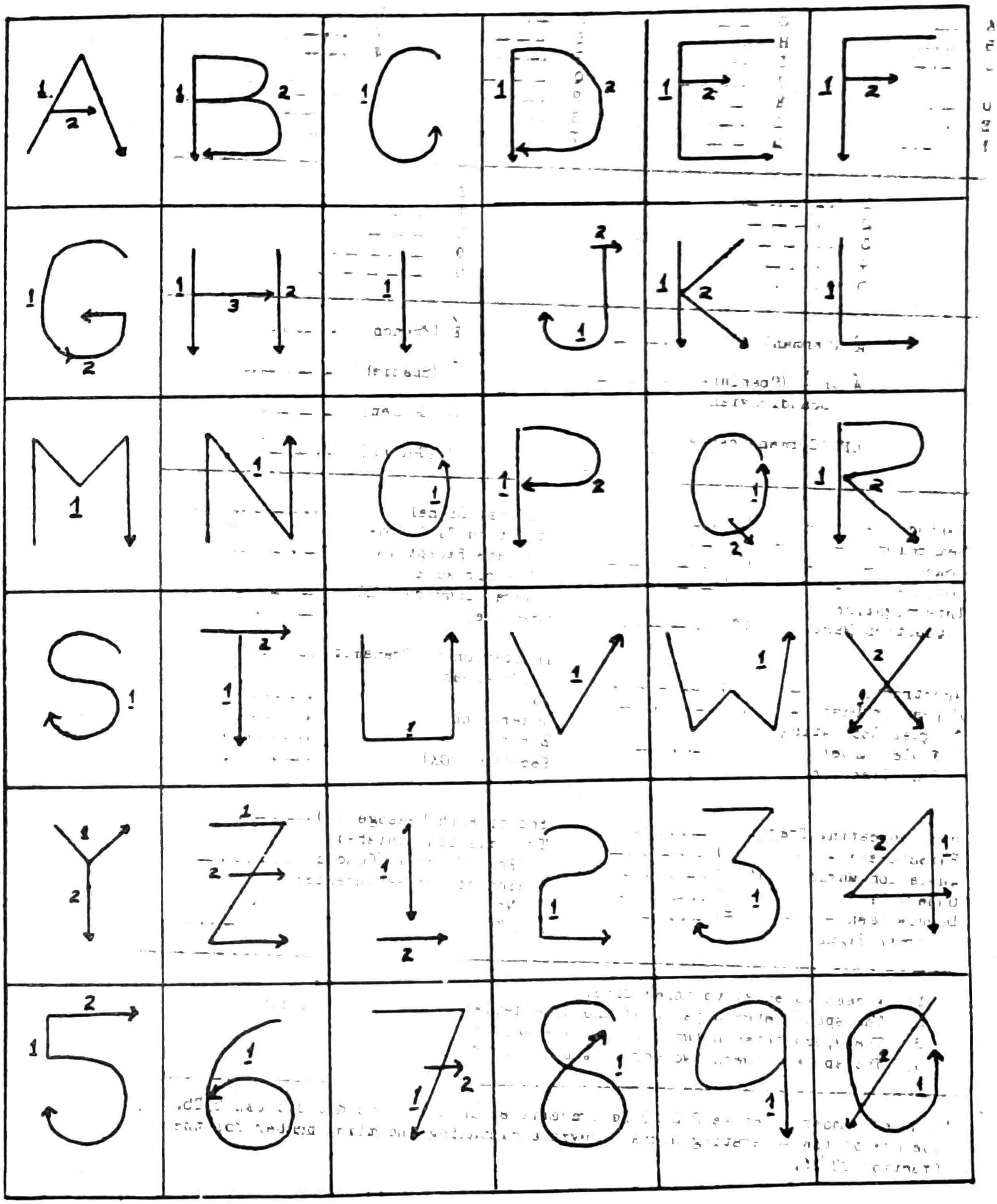


Table 4

A. PHONETIC ALPHABET: as used by the Army and Navy, and by all the United Nations, especially in communications over the telephone, to spell out proper names and the code and cipher texts of messages, effective February 1, 1943.

Letter	Spoken as	Letter	Spoken as	Letter	Spoken as
A	ABLE	J	JIG	S	SUGAR
B	BAKER	K	KING	T	TARE
C	CHARLIE	L	LOVE	U	UNCLE
D	DOG	M	MIKE	V	VICTOR
E	EASY	N	NAN	W	WILLIAM
F	FOX	O	OBOE	X	X-RAY
G	GEORGE	P	PETER	Y	YOKE
H	HOW	Q	QUEEN	Z	ZEBRA
I	ITEM	R	ROGER		

NUMERALS

Numeral	Spoken as	Numeral	Spoken as
Ø	ZE-RO	5	FI-YIV
1	WUN	6	SIKS
2	TOO	7	SEV-VEN
3	THUH-REE	8	ATE
4	FO-WER	9	NI-YEN

Examples:

MARY - MIKE ABLE ROGER YOKE.

LYON - LOVE YOKE OBOE NAN.

JIXPO - JIG ITEM XRAY PETER OBOE.

23 - TOO THUH-REE.

5Ø - FI-YIV ZE-RO.

4ØØ - FO-WER HUNDRED.

1627 - WUN SIKS TOO SEV-VEN.

18000 - WUN ATE THOUSAND.

B. PHONETIC ALPHABET: as used by the cable and radio companies.

Letter	Spoken as	Letter	Spoken as	Letter	Spoken as
A	ARTHUR	J	JOHN	S	SAMUEL
B	BENJAMIN	K	KITTY	T	THOMAS
C	CHARLIE	L	LOUIS	U	UNION
D	DAVID	M	MARY	V	VICTOR
E	EDWARD	N	NATHAN	W	WILLIAM
F	FRANK	O	OTTO	X	X-RAY
G	GEORGE	P	PETER	Y	YELLOW
H	HENRY	Q	QUEEN	Z	ZERO
I	ISAAC	R	ROBERT		

BOOKS IN THE CRYPTOGRAPHIC SERIES

C-1 MANUAL FOR THE SOLUTION OF MILITARY CIPHERS, Parker Hitt
C-2 CRYPTANALYSIS OF THE SIMPLE SUBSTITUTION CIPHER, Wayne G. Barker
C-3 ELEMENTS OF CRYTANALYSIS, William F. Friedman
C-4 STATISTICAL METHODS IN CRYPTANALYSIS, Solomon Kullback
C-5 CRYPTOGRAPHY AND CRYPTANALYSIS ARTICLES, Vol 1, Friedman
C-6 CRYPTOGRAPHY AND CRYPTANALYSIS ARTICLES, Vol 2, Friedman
C-7 ELEMENTARY MILITARY CRYPTOGRAPHY, William F. Friedman
C-8 ADVANCED MILITARY CRYPTOGRAPHY, William F. Friedman
C-11 SOLVING GERMAN CODES IN WORLD WAR I, William F. Friedman
C-14 MANUAL OF CRYPTOGRAPHY, Luigi Sacco
C-17 CRYPTANALYSIS OF THE HAGELIN CRYPTOGRAPH, Wayne G. Barker
C-20 HISTORY OF CODES AND CIPHERS IN THE U.S. PRIOR TO WORLD WAR I, ed. Barker
C-21 HISTORY OF CODES AND CIPHERS IN THE U.S. DURING WORLD WAR I, ed. Barker
C-22 HISTORY OF CODES AND CIPHERS IN THE U.S. DURING THE PERIOD BETWEEN THE WORLD WARS, PART I, 1919-1929, ed. Barker
C-26 CRYPTANALYSIS OF AN ENCIPHERED CODE PROBLEM, Wayne G. Barker
C-30 MILITARY CRYPTANALYSIS, PART I, William F. Friedman
C-33 COURSE IN CRYPTANALYSIS, Volume 1, British War Office
C-34 COURSE IN CRYPTANALYSIS, Volume 2, British War Office
C-35 THE ORIGIN AND DEVELOPMENT OF THE NATIONAL SECURITY AGENCY, George A. Brownell
C-36 TREATISE ON CRYPTOGRAPHY, André Lange and S.A. Soudart
C-39 CRYPTANALYSIS OF SHIFT-REGISTER GENERATED STREAM CIPHER SYSTEMS, Barker
C-40 MILITARY CRYPTANALYSIS, PART II, William F. Friedman
C-41 ELEMENTARY COURSE IN PROBABILITY FOR THE CRYPTANALYST, Andrew M. Gleason
C-42 MILITARY CRYPTANALYTICS, PART I, VOL. 1, Friedman & Callimahos
C-43 MILITARY CRYPTANALYTICS, PART I, VOL. 2, Friedman & Callimahos
C-44 MILITARY CRYPTANALYTICS, PART II, VOL. 1, Callimahos & Friedman
C-45 MILITARY CRYPTANALYTICS, PART II, VOL. 2, Callimahos & Friedman
C-46 PATTERN WORDS — THREE LETTERS TO EIGHT LETTERS IN LENGTH, Carlisle
C-48 PATTERN WORDS — NINE LETTERS IN LENGTH, Sheila Carlisle
C-49 THE INDEX OF COINCIDENCE AND ITS APPLICATIONS IN CRYPTANALYSIS, Friedman
C-50 CRYPTOGRAPHIC SIGNIFICANCE OF THE KNAPSACK PROBLEM, O'Conner and Seberry
C-52 THE AMERICAN BLACK CHAMBER, Herbert O. Yardley
C-53 TRAFFIC ANALYSIS AND THE ZENDIAN PROBLEM, L.D. Callimahos
C-54 HISTORY OF CODES AND CIPHERS IN THE U.S. DURING THE PERIOD BETWEEN THE WORLD WARS, PART II, 1930-1939, ed. Barker
C-55 INTRODUCTION TO THE ANALYSIS OF THE DATA ENCRYPTION STANDARD (DES), Barker
C-56 ELEMENTARY CRYPTOGRAPHY AND CRYPTANALYSIS, Donald D. Millikin
C-57 SECRETS CIPHERS OF THE 1876 PRESIDENTIAL ELECTION, D. Beaird Glover
C-58 SOLVING CIPHER PROBLEMS, Frank W. Lewis
C-59 CRYPTANALYSIS OF THE SINGLE COLUMNAR TRANSPOSITION CIPHER, W. Barker
C-60 MILITARY CRYPTANALYSIS, PART III, William F. Friedman
C-61 MILITARY CRYPTANALYSIS, PART IV, William F. Friedman
C-62 PATTERN WORDS — TEN-LETTERS AND ELEVEN-LETTERS IN LENGTH, Wallace
C-63 PATTERN WORDS — TWELVE-LETTERS AND GREATER IN LENGTH, Wallace
C-64 U.S. NAVAL CRYPTOGRAPHIC ACTIVITIES IN THE PHILIPPINES PRIOR TO WWII, ed. Carlisle
C-65 U.S. NAVAL COMMUNICATIONS INTELLIGENC

ISBN: 0-89412-173-1

מבללת אורט בראודה
תשפ"ה
003351602046
ELEMENTARY CRYPTOGRAPHY A